Available online at www.sciencedirect.com

SCIENCE DIRECT

Developmental Biology 279 (2005) 271–290

DEVELOPMENTAL BIOLOGY

www.elsevier.com/locate/ydbio

Mouse ovary developmental RNA and protein markers from gene expression profiling

Luisa Herrera[a,1,2], Chris Ottolenghi[a,b,1], J. Elias Garcia-Ortiz[a], Massimo Pellegrini[b], Francesca Manini[b], Minoru S.H. Ko[a], Ramaiah Nagaraja[a], Antonino Forabosco[b], David Schlessinger[a,*]

[a]*Laboratory of Genetics, Gerentalogy Research Centre, National Institute on Aging, Suite 3000, 333 Cassell Drive, Baltimore, MD 21224, USA*
[b]*Medical Genetics, Department of Mother and Child, University of Modena and Reggio-Emilia, Policlinico, Via del Pozzo 71, 41100 Modena, Italy*

Received for publication 15 October 2004, accepted 17 November 2004
Available online 12 January 2005

Abstract

To identify genes involved in morphogenetic events during mouse ovary development, we started with microarray analyses of whole organ RNA. Transcripts for 60% of the 15,000 gene NIA panel were detected, and about 2000 were differentially expressed in nascent newborn compared to adult ovary. Highly differentially expressed transcripts included noncoding RNAs and newly detected genes involved in transcription regulation and signal transduction. The phased pattern of newborn mouse ovary differentiation allowed us to (1) extend information on activity and stage specificity of cell type-specific genes; and (2) generate a list of candidate genes involved in primordial follicle formation, including podocalyxin (*Podxl*), PDGFR-β, and a follistatin-domain-encoding gene *Flst1*. Oocyte-specific transcripts included many (e.g., *Deltex2*, *Bicd2*, and *Zfp37*) enriched in growing oocytes, as well as a novel family of untranslated RNA's (RLTR10) that is selectively expressed in early stage follicles. The results indicate that global expression profiling of whole organ RNA provides sensitive first-line information about ovarian histogenesis for which no in vitro cell models are currently available.
Published by Elsevier Inc.

Keywords: Microarray; Ovary; Primordial follicles; Development

Introduction

Mammalian ovary development coordinates complex molecular, cellular, and histogenetic events. Contrary to testis differentiation, where germ cells are dispensable, oocytes are required to induce ovarian somatic cells (McLaren, 2000). In turn, the accretion of somatic cells around the oocytes forms primordial follicles that preserve the oocyte pool (Ohno and Smith, 1964; Peters, 1969). Subsequently, primordial follicles are recruited to grow. Initially controlled by local intercellular signaling, they later respond to additional regulation of selective ovulation by the hypothalamic–pituitary axis.

Experimental embryology has inferred that the perinatal formation of follicles and the reorganization of the ovary into morphological compartments ("cortex" and "medulla") are required to prevent the otherwise massive derepressed growth of ovarian follicles (Byskov et al., 1997 and references therein). Consistent with these findings, we have shown that in mice lacking the *Foxl2* transcription factor, deregulated oocyte growth (Schmidt et al., 2004) follows a primary impairment of follicle formation and ovarian histogenesis (Uda et al., 2004).

In the mouse model, numerous other factors have been identified that directly affect primordial follicle formation (*Fig-α*, Liang et al., 1997; Soyal et al., 2000; Dean, 2002; *Wnt4*, Vainio et al., 1999; *TrkB*, Spears et al., 2003; *NGF*, Dissen et al., 2001) or the related progression into prophase of meiosis I (*Dazl*, Ruggiu et al., 1997). Other genes are required earlier for the establishment of the germ cell pool (*c-kit/Steel*, *Zfx*, Elvin and Matzuk, 1998), or at later stages

* Corresponding author. Fax: +1 410 558 8331.
E-mail address: schlessingerd@grc.nia.nih.gov (D. Schlessinger).
[1] These authors contributed equally.
[2] Current address: Programa de Genética Humana, ICBM, Facultad de Medicina, Universidad de Chile, Santiago, Chile.

0012-1606/$ - see front matter. Published by Elsevier Inc.
doi:10.1016/j.ydbio.2004.11.029

for follicle recruitment (*Foxo3*, Castrillon et al., 2003) or growth. Growth is regulated locally (*Gdf9*, *Lif*, and *Bmp15*; Elvin et al., 1999; Yan et al., 2001) or in response to pituitary endocrine factors (*Fshr*, *Lhr*, Elvin and Matzuk, 1998). Still other factors promote growth (e.g., cyclin D2) or regulate apoptosis (e.g., *caspase-9* and-*2*, and *bcl-2* gene family members; reviewed by Pru and Tilly, 2001; Reynaud and Driancourt, 2000).

To help assess the full range of genes involved in follicle development, we have turned to a global genomic approach. Nascent primordial follicles occupy most of the organ at birth, but constitute only a marginal fraction of the adult ovary. This makes parallel gene expression profiling on whole organs with microarrays of cDNAs a useful way to infer genes that are active specifically during primordial follicle formation. A comparable whole organ approach was recently used to study stages of spermatogenesis (Fujii et al., 2002) and molecular anomalies in ovaries of mice lacking Fsh-β (Burns et al., 2001). Here we infer that at least 9000 transcripts were expressed during mouse ovary formation. Many showed developmentally regulated transcription. In addition to transcripts already known to be involved in ovary development, they included novel untranslated RNAs and mRNAs for signaling and transcription factors that had not previously been detected in the ovary. We used further experiments and complementary information to assign cell and stage specificity to selected transcripts.

Methods

Probe preparation and microarray hybridization

Animals were euthanized ethically according to ACUC-approved NIA Animal Protocols. Total RNA samples from C57BL/6J mice were used for microarray hybridization and quantitative real-time RT-PCR. Mouse tissues were sampled and immediately rinsed with phosphate-buffered saline (PBS) and frozen on dry ice. Trizol (Life Technologies) was used to isolate total RNA. For real-time PCR, the RNA samples were treated with "RNase-free DNase" (Boehringer Mannheim, Mannheim, Germany) to eliminate genomic DNA contamination. Total RNAs from adult ovary (4–6 months) were used in real-time RT-PCR experiments. The mouse NIA 15K gene set (15,264 rearrayed unique genes or gene isoforms), printed on seven nylon membranes, was used in hybridizations (Tanaka et al., 2000). Sequence information is available at http://www.ncbi.nlm.nih.gov/ and at the National Institute on Aging web site http://lgsun.grc.nia.nih.gov/. The cDNA clone identification, for example, H3146E11, is the ID code used in the NIA 15k gene set (Tanaka et al., 2000).

Three cDNA samples were prepared for adult (4–6 months old) and newborn ovary (2–3 days postnatum). The template for cDNA synthesis was 3.4 μg of total RNA using SuperScript II reverse transcriptase with oligo(dT) primer. Of the amplified cDNA, 2.1 μg was used as substrate to incorporate [α-^{33}P]dCTP (Amersham Pharmacia Biotech Inc, Piscataway, NJ, USA) by random priming (RadPrime DNA labeling system, Life Technologies). The DNA probes were then purified on Quick-spin Sephadex G25 columns (Roche Diagnostics). Prehybridization and hybridization were done as in Tanaka et al. (2000). Results from three independent hybridizations were obtained for each probe. Images were analyzed by Imagequant 5.0 (Amersham Pharmacia).

Analysis of data

We used several current analytic methods for the analysis of gene expression profiling data, focusing primarily on genes differentially expressed in newborn compared to adult ovary.

Normalization and analysis of expressed and differentially expressed genes

Statistical analysis of microarray experiments is much affected by the preliminary step of normalization of data, which is required to remove systematic variations in signal intensity (e.g., Park et al., 2003). We used several standard approaches to optimize the sensitivity and specificity of inferred lists of expressed genes, as follows. First, background was subtracted by linear interpolation of local values determined at 432 widely distributed microarray filter locations that bore no DNA. Similar to previous analyses based on the same arrays (Cui et al., 2002), the intensities were sum normalized and a threshold value of 100,000 arbitrary units on average across replicates was used as an inclusion criterion. (This corresponds to 1.5 times average background after normalization, a level at which signals are clearly visible by eye on microarray images.) In addition, the lowest values corresponding to 4% of the total (3247 of 91,488) were replaced by surrogate values using a conservative method ("average row imputer", as in the option of the SAM program; Tusher et al., 2001, and http://www-stat.stanford.edu/~tibs/SAM/).

Next, to normalize sample variance across experiments, we separately applied two standard methods, both available in the Focus program (Cole et al., 2003), to the trimmed, sum-normalized data set. In one, a modified Z transformation assigned to all replicates the same mean as the mean value and the same standard deviation (SD) as the mean SD across all data sets. This preserves a metric, based on intensities, that is similar to the original data. In a second method, a tool to deal with overall data variability (Sidorov et al., 2002), rank normalization, was used to assign the rank of each gene (from 1 to n) as the ranked value for that gene in each replicate. This approach emphasizes relative expression levels rather than actual intensities. (We also tried direct normalizations by global scaling approaches (Kroll and Wolfl, 2002), but they gave less sensitivity for the recovery of known positive clones than the other

methods (data not shown).) The two lists obtained overlap extensively, but each added some clones to a merged list (see below). The resulting data set is given in WEB Table 1. Based on amplified cDNA probes from newborn and adult ovary, 9582 clones (~60% of the total) showed mean expression intensities above background in newborn ovary, adult ovary, or both.

Genes differentially expressed in newborn or adult ovary were inferred from the total list with the Focus implementation of a PRIM algorithm for microarray analysis (Cole et al., 2003). The approach mimics human analysts' inferences about reliability of expression differences, employing both raw intensities and ratios of values controlled by an explicit statistical criterion. Selection of differentially expressed genes was based on Boolean rules of minimal fold ratio and mean raw unit differences, optimized for 60% reliability in predicting 1.5 mean fold changes in normalized data sets. This is equivalent to setting a 40% false discovery rate (FDR, as in, e.g., Storey and Tibshirani, 2003), a standard estimate of the proportion of false positives expected from repeated experiments. The validation we performed (see below) indicates that this FDR value actually reflected a lower fraction of false positives (20%), as calculated from the redundant clones in our data set. This is consistent with reported simulations (Cole et al., 2003). The hits in WEB Tables 2 and 3 are ordered according to their score (a measure of the normalized distance from the mean for four parameters of interest) as calculated by Focus.

In simulations, Focus outperformed other commonly used analytical tools (e.g., Tusher et al., 2001) in its ability to detect higher numbers of known true positives while maintaining comparably high specificity (Cole et al., 2003). For the sets of genes in WEB Table 1, we selected significant genes as those inferred with the modified Z transformation because of its very high specificity and sensitivity. That group was augmented 19% by merging it with the list recovered by rank normalization, thereby including many redundant clones with consistent expression, as follows.

Critical controls for internal consistency were provided by transcripts redundantly represented on microarray membranes. Based on NIA annotations, 1534 mouse UniGene clusters are represented by at least two clones in the 15K microarray, amounting to 23% of the total. The transcripts with at least two clones among those called significant were classified as consistent (those where all clones had been assigned to one ovary stage) or inconsistent (those with assignments to both stages). Using this approach (Table 1), with z transformation, we obtained 97% specificity (equivalent to 3% FDR) and 57% sensitivity (a total of 1582 differentially expressed clones, 471 of them in newborn ovary). With rank normalization, 74% specificity was obtained, with 35% sensitivity (a total of 1316 clones, 323 in the newborn). Merging the results from the two approaches (see above) yielded 80% specificity and 50% sensitivity overall. The high specificity for this large number of significant calls supports the quality of the merged set, with

Table 1
Specificity and sensitivity of inferred differentially expressed genes

Samples	Normalization	Specificity
Merged (Nb + Ad)	Modified-Z	66/70 (97%)
	Rank	40/54 (74%)
	Merged	83/105 (**79**%)
Samples	**Normalization**	**Sensitivity**
Newborn (Nb)	Rank	227/538 (42%)
	Modified-Z	93/538 (17%)
Adult (Ad)	Rank	173/629 (27%)
	Modified-Z	320/629 (51%)
Merged (Nb + Ad)	Rank	400/1167 (34%)
	Modified-Z	413/1167 (35%)
	Merged	600/1167 (**51**%)

Internal consistency of the lists of differentially expressed genes based on redundant UniGene clusters in the NIA 15K data set (see Methods). Calculations were done for each of two normalization methods for newborn and adult (Focus software, see Methods), and for a merged set of clones. The merged set produced the total of 2247 clones differentially expressed (Web Tables B and C, and see text). Upper panel: specificity is calculated as the fraction of UniGene clusters with clones assigned to the same developmental stage, among the total number of UniGene clusters with at least two clones being differentially expressed. Lower panel: sensitivity is calculated as the number of clones that have been detected divided by the total number of clones that belong to detected UniGene clusters.

an inferred total of 2247 differentially expressed transcripts (assembled into 2115 UniGene clusters). Of these, 889 (850 UniGene clusters) are more highly expressed in newborn ovary. Fig. 1 shows overall distributions. The resulting full lists of genes are in WEB Tables 2 and 3, and subsets of selected genes are in Tables 2–4 (see Results).

Validation of microarray expression profiling

In addition to Northern blot hybridization (Results, Fig. 2) and agreement of values for redundant clones, high-level expression of genes was preliminarily validated by other available data. The NIA 15K array was recently used to compare ovary expression profiles of normal and FSH-β null mice, detecting 13 differentially expressed genes (Burns et al., 2001). Because Fsh is required from the preantral follicle stage to ovulation, the differentially expressed genes should be upregulated in adult ovaries; and consistent with that notion, six of these genes were found differentially expressed in our analysis, all of them in adult ovary (Star, Hsd3b1, Cyp11a, Ldh2, Tagln, Rbbp7). Similarly, adult ovary showed enrichment for 11 of 27 genes differentially expressed in granulosa cell tumors associated with loss of function of Inha (Burns et al., 2001). This is consistent with their direct or indirect role in granulosa cell proliferation, which is suspended neonatally during the process of follicle formation (e.g., Rajah et al., 1992). Only 1 of the 27 genes (damage-specific DNA binding protein, *Ddb1*) was scored as upregulated in our newborn ovary sample.

Gene ontology and INTERPRO analysis

For global functional classification, a publicly available package to map UniGene clusters to the Gene Ontology

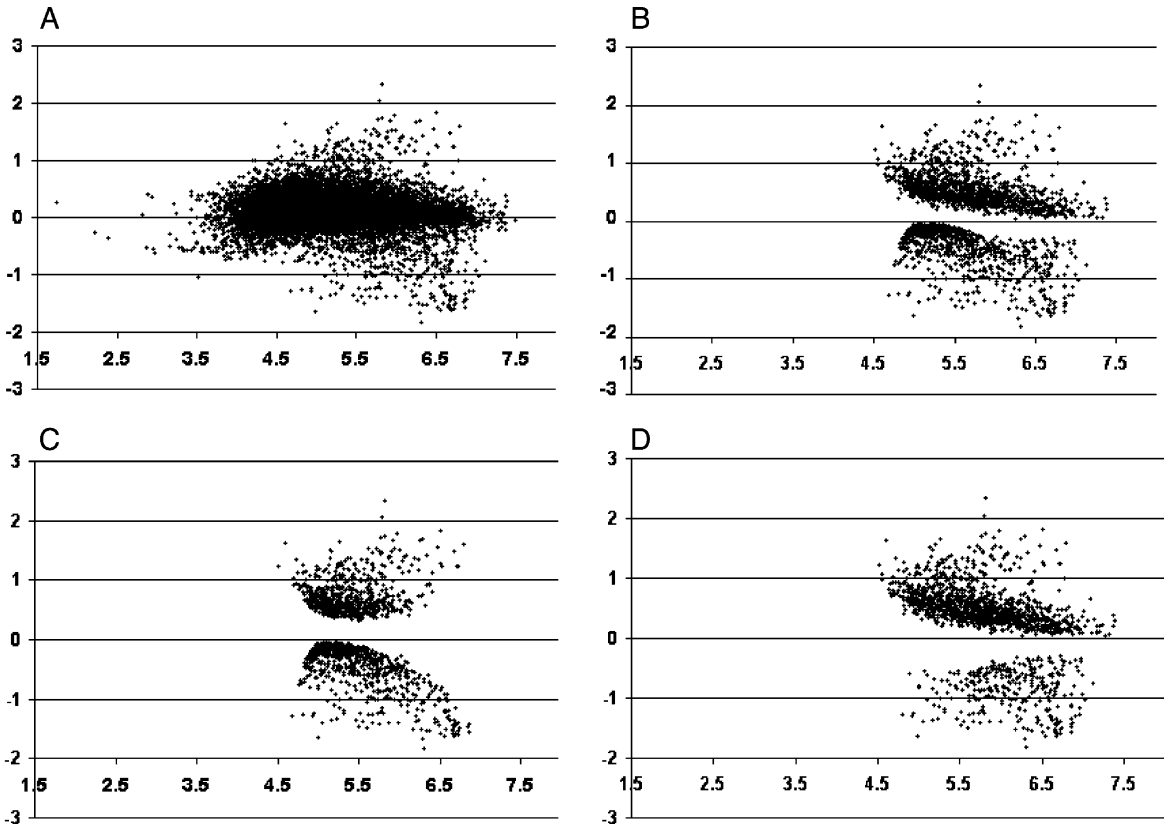

Fig. 1. Scatter plot of mean intensity adult/newborn log ratio (y-axis) vs. log of the geometric mean intensity (x-axis) for the entire data set (A) and for the significant hits detected using two normalization methods (modified Z and rank normalization, C and D, respectively, as well as the merged set of significant hits, B; see Methods).

database (MAPPfinder version 2.0) inferred the most abundant entries in either newborn or adult ovary according to standard Gene Ontology classification (Fig. 3A). In a complementary approach, protein domain information was derived from the NIA GeneIndex (http://lgsun.grc.nia.nih.gov/geneindex/) and a combined search through INTER-

Table 2
Functional classification of inferred differentially expressed genes

Annotated genes	Newborn	Adult
	Examples	
Apoptosis	Casp3, Zac1, Tia1, F830014G06 (Riken)	Gpx1, Bag3–4, Gadd45g, Tnfrsf12a, Bcl2l10, Cul1, Mcl1, Sgk, Tpt1, App, scotin, Trp53, Dad1, Ngfrap1
Immune system	–	C1qr1, C1qb, MyD88, ID_B1, B2_Mm2, H2-Eb1
Oxidative stress	–	Gpx1, Prdx1, Sod1
Steroidogenesis	–	Cyp11A1/P450ssc, Cyp17, StAR, Hsd3b
Transcription	Sox4, Sox15, Nobox/Og2x, Psx1, Pknox1/Prep1, Meis1, Idb1, Trim28, Tceb1l, Mcmd5, Hmga1, Hmgb3, Polr2a, C76800, Ankrd6, Preb, Taf7, Taf7L, Mllt10, Polr2j, Ldb1, Phtf, 1500004O14, 9830141C09(Riken), Supt4h, Gtf2h1, Ubp1, Stat3, Psmd10, Cnbp, Gtf2b, Stat3, Pole4, Ccnc, Hdac2, Dnmt3a, Nr2f2, Hmga2, Pole4, E2f6, Mycbp, 5730434I03(Riken), Lztr1, Usf1, Hdac3	SF1, Gli2, Hoxc9, Onecut1, Gata2, Fosl1, -2, Heyl, 2610103K11Rik, Fhl2, Sh3d3, Rabl3, Irf1, Runx2, D15Ertd417e, Atf3, Nfe2l2, A730098D12(Riken), Tef, Polr2c, Copeb, Nrf1, Tgfb1i4, Foxo1, Tgfb1i4, Myc, Baz1a, Tgfb1i4, Rpo1-1, Cbfa2t1h, Paf53, Ankrd106, Hif1a, Hmgb2, Tcf4, Atf2, Cnot8, Rab7, Tcfe2a, Aes, Trp53, Mcmd
RNA binding proteins	1200009K13(Riken), Tarbp2, Raly, D11Ertd619e, Sfrs3, G3bp2, 1200009K13Rik, Supt6h, Zfr, Eif4e, U2af1, Eif4g2, Slbp, Fmr1, Ps1d, Tia1, Dazl, Stau1, Ssfa1	Hnrpdl, Srp19, C81487, 2310046H11Rik, Snrpb2, G3bp, Srp9, Hnrpa1, Ddx1, -5, Rps14, Rps27, Sfrs2, Park7
Secreted factors	Npnt, Podx1, Ag2R2, Wnt4, Wnt2, Semaphorins-6a, -4g, Itgb7, Ptn, P4hb, Igf2, Igfbp5, Spon2, Prss15, Abcd3, Tem8, Cd97, Serpina6, Akp5, Acadl, Timd2, Entpd2, Tparl, Mesp2	Zp3, Pecam, Sel1, Eg1, Gdf9, Bmp15, TGFbeta2 Ctsl, Robo1, Ctsb, Itgb5, Cst3, Gm2a, Sdf4, Gpx3 Serpine2, Frk, Dpp7, Cdh1, Scg3, Cd8b, Serpinh1, Inha, Ecm1, C1qb, C1qr1, Plxna3, Lrp1, Col4a1, -2 Rassf5, Jak1
Cytoplasmic factors	Apc, Nras, Ralgds, Rtn4, CK18, CK19, CK8	

Functional classification of selected differentially expressed genes classified into structural (transcription, RNA binding, cytoplasmic, secreted) and functional (apoptosis, immune system, oxidative stress, steroidogenesis) categories according to annotations from Gene Ontology, Riken, and NIA databases.

Table 3
Genes preferentially expressed in oocytes, inferred from microarray data as in WEB Tables 2 and 3 combined with information from EST libraries (see Methods)

NIA clone ID	ESTs			Gene/UniGene/Riken	Annotations
	Egg	*NbOv*	*p*		
A. Inferred oocyte-enriched genes expressed throughout oogenesis					
H3028B04	58	0	0	*1700022N24*	IPR001841 Zn-finger, RING, C3HC4
H3092A01	53	0	0	*Chek1*	checkpoint kinase 1 homolog (*S. pombe*)
H3071A11	50	1	0	*G3bp2-pending*	Ras-GTPase-activating protein SH3-binding rotein 2; RRM domain
H3014B10	36	4	0	*Ncl*	nucleolin
H3113A01	28	1	0	*Slbp*	stem loop binding protein
H3075B10	24	0	0	*2810404F18*	PAK-box/P21-Rho-binding; non-kinase CDC42 effector
H3102E07	17	0	0	*Mater*	maternal effect gene
H3074F07	20	1	0.001	*Usp7*	Ubiquitin thiolesterase, family 2; IPR002083
H3059G02	14	0	0.001	*Mm.222875*	Intron-containing, no long open reading frame, no similarity by BlastX
H3076A09	13	0	0.002	*Mllt10*	myeloid/lymphoid or mixed lineage-leukemia translocation 10
H3152D10	13	0	0.002	*0610039L19*	dTDP-4-dehydrorhamnose reductase
H3076G12	10	0	0.009	*C430041I18*	Similar to hypothetical protein MGC4039
H3032A06	8	0	0.022	*Pabpc4*	RRM and polyA-binding
H3074A05	14	2	0.027	*Prc1*	protein regulator of cytokinesis 1-like (LOC233406)
H3091D12	7	0	0.036	*Golph2*	Mak10 subunit, NatC N/alpha-terminal acetyltransferase
H3093E01	7	0	0.036	*Rtn4*	reticulon 4/Nogo (also found in somatic tissues)
H3140E07	7	0	0.036	*2410017P07*	Uncharacterized "DUF1167" domain encoding protein
H3121D01	9	1	0.061	*Pole4*	DNA polymerase epsilon 4 (p12 subunit) (also, H3126A12)
H3125E11	9	1	0.061	*P4hb*	disulfide isomerase
H3140F04	2	5	0.077	*Anxa6*	annexin A6 (Anxa6)
H3005C12	7	1	0.13	*Trim28*	tripartite motif protein 28; TIF1-beta; KRIP1
H3140G02	7	1	0.13	*Pes1*	pescadillo homolog 1, containing BRCT domain
H3123C01	2	4	0.151	*D0H4S114*	DNA segment, human D4S114
H3130A05	1	3	0.156	*Psmd2*	proteasome 26S subunit, non-ATPase, 2 (Psmd2)
H3134F02	1	3	0.156	*Mm.15755*	A130024G22 Riken: transcriptional regulator, SIN3A/yeast
H3136A02	6	1	0.188	*Sh3d1B*	SH3 domain protein 1B
H3071G01	5	6	0.202	*AA407558/Mm.245522*	Ankyrin repeat domain protein 17
H3081F02	8	2	0.203	*Eif4enif1*	translation initiation factor 4E nuclear import factor 1
H3157H12	8	2	0.203	*Rock1*	Rho-associated coiled-coil forming kinase 1
H3149E08	13	5	0.265	*Anxa7*	annexin A7 (Anxa7)
H3030C03	2	3	0.282	*2610511G16*	IPR003034 DNA-binding SAP
H3140F06	9	8	0.294	*P5-pending*	disulfide isomerase A6 P
H3081E09	1	2	0.322	*Tp120a-pending*	TBP-interacting 120
H3147B01	1	2	0.322	*Tyro3*	protein tyrosine kinase 3
H3137B11	4	4	0.358	*Mest*	mesoderm specific transcript
H3073F11	6	2	0.361	*Simp-pending*	1300006C19 Riken; IPR003674 Oligosaccharyl transferase
H3117B10	6	2	0.361	*2310046H11*	RRM-encoding
H3036G12	4	1	0.374	*Krt2–8 ?*	PR000719 Protein kinase; IPR003054 Type II keratin;
H3081C01	3	3	0.412	*2810411G23*	2810411G23Riken, similar to Tumor protein D52
H3140E04	2	2	0.488	*Hcph*	hemopoietic cell phosphatase
H3033A12	4	3	0.534	*1110061O04 en*	tyrosine specific and dual specificity phosphatase
H3119C06	1	1	0.614	*Nol5*	nucleolar protein 5
H3131B12	1	1	0.614	*5730434I03*	BTF3 transcription factor
H3142G03	1	1	0.614	*Ddb1/Dnajc3*	DnaJ (Hsp40) homolog, subfamily C, member 3
H3144A01	1	1	0.614	*Mm.103143/325551*	Alanine-rich region/Type I antifreeze protein containing protein
H3146G11	1	1	0.614	*MGC11792*	Similar to Inversin
H3157E01	1	1	0.614	*D030016E14*	OVARC1001768
H3024G11	1	1	0.614	*4833420O05*	hypothetical Ubiquitin carboxyl-terminal hydrolase family 2
H3024E01	2	1	0.678	*MGC19382*	PR001026; Phosphatidylinositol binding clathrin assembly
H3130E06	2	1	0.678	*H19*	Untranslated RNA (also H3130H06,H3140G12)
H3134D02	2	1	0.678	*2610511O17*	2610511O17 Riken
H3076B02	2	1	0.678	*Zfr*	zinc finger RNA binding protein
H3145H09	2	1	0.678	*Prdx4*	peroxiredoxin 4 (Prdx4)
B. Inferred oocyte-enriched genes expressed in growing oocytes					
H3056F01	178	0	0	*Gdf9*	Growth differentiation factor 9 (also H3053D04)
H3079C07	92	0	0	*Rhpn2*	GTP RHO binding protein
H3031C11	35	0	0	*Myo10*	Myosin X

(continued on next page)

Table 3 (*continued*)

NIA clone ID	ESTs			Gene/UniGene/Riken	Annotations
	Egg	*NbOv*	*p*		
B. Inferred oocyte-enriched genes expressed in growing oocytes					
H3004E12	34	0	0	*Dtx2*	Deltex2 (Notch pathway)
H3068G09	31	0	0	*Spin*	Maternal transcript Spindlin
H3065G04	27	0	0	*Prostein-pending*	Rieske iron-sulfur protein (also H3066G04)
H3146F12	26	2	0	*2700023B17*	C19orf13 protein (AlphaSNBP(B))
H3052A09	24	0	0	*E330034G19*	Aldehyde dehydrogenase-domain
H3042G08	22	0	0	*Gpr108*	Lung seven transmembrane receptor 2
H3056E04	21	0	0	*4921517A06*	IPR007527 SWIM Zn-finger
H3122F01	19	0	0	*Atp5b*	ATP synthase, H+ mitochondrial F1 complex
H3056E09	17	0	0	*Zfp37*	KRAB-box zinc finger
H3023D06	15	1	0.005	*Pgk1*	phosphoglycerate kinase 1
H3103E12	12	0	0.003	*Bcl2l10*	Bcl2-like 10 (also H3058C03)
H3070C09	9	0	0.014	*Mm.222875/Mm.333241*	Intron-containing; no ORF
H3054E08	9	0	0.014	*D1Ertd251e*	2900026B15 Riken
H3052D01	9	0	0.014	*6330415M09*	hypothetical Doublecortin (also H3055A02)
H3068E12	8	0	0.022	*Cul1*	SCF E3 ubiquitin ligase complex
H3066G06	8	0	0.022	*LOC234699*	Human autoantigen Ge-1
H3056G11	8	0	0.022	*Nub1*	NEDD8 ultimate buster-1 (BS4 protein)
H3055A10	8	0	0.022	*Tef*	Cytosolic phospholipase A2
H3006A06	8	0	0.022	*Surf4*	Surfeit gene 4
H3055E09	8	0	0.022	*E130301L11*	Tastin/Trophinin assisting
H3129A03	11	1	0.027	*2410073G16*	Similar to KIAA0261 [Homo sapiens]
H3122D01	7	0	0.036	*Nedd4*	Ubiquitin-protein ligase
H3109D02	7	0	0.036	*4930504E06*	NF-E2 inducible protein
H3054G09	7	0	0.036	*Banp*	Btg3 associated nuclear protein

(A) Newborn enriched: expressed at high levels in ovulated eggs but also showing higher relative expression in newborn ovary microarray because of the dilution of oocytes in adult ovary. (B) Adult enriched: growing oocyte markers (expression is higher in adult ovary microarray and in egg EST library than newborn microarray and EST library, respectively). These genes were determined by combining EST library information (unfertilized egg and newborn ovary) with microarray data. The numbers of ESTs detected and *P* value for the comparison between EST libraries are given.

PRO and RIKEN databases (at http://fantom2.gsc.riken.go.jp). The 15K clone sequences were mapped onto the RIKEN database (60,770 sequence assemblies) under stringent criteria: matches with high sequence identity over regions compatible with the length and sequence errors typically detected in EST sequences, that is, 95% identity over 40 nucleotides. The RIKEN annotations provide links to the encoded protein sequences according to the Trembl nomenclature (fantom2.gsc.riken.go.jp, Carninci et al., 2003), and the latter is separately linked to INTERPRO protein domain annotations (ftp.ebi.ac.uk/pub/databases/interpro). A data set assembling this information was created, and the top-ranking protein domains were sorted for both the newborn and the adult ovary samples (Fig. 3B).

EST analysis

The sequences derived from the 15K NIA and RIKEN databases (see previous section) were mapped onto two mouse EST libraries, from newborn ovary (7773 sequences, NIA; Tanaka et al., 2000) and fertilized oocytes (library #10029; 12,736 sequences, NIA; Sharov et al., 2003) (all sequences available at www.ncbi.nlm.nih.gov/UniGene). The relative abundance of the EST sequences and the associated statistical differences were calculated (Audic and Claverie, 1997); the differentially expressed hits compared to the microarray data for equivalent tissues are given in Table 3.

Public microarray data

Microarray data on RNA expression levels across a wide range of adult mouse tissues are publicly available (Su et al., 2002). They include ovary, uterus, oocytes, and fertilized eggs. We reasoned that some genes with high expression in somatic cells of the ovary could be recognized by a high signal intensity in both ovary and uterus combined with low expression in two female germ-line samples, referred to global median intensity across all tissues tested (available at http://symatlas.gnf.org/SymAtlas/); 314 genes in the NIA 15K set satisfied the criteria. About 22% of them (70/314) were differentially expressed between newborn and adult ovary (Table 4), as expected from the overall representation of differentially expressed genes (2247/9582). This list provided candidates for somatic lineage transcripts (see Results).

Analysis of RLTR10 repeat family

The H3146E11 transcript sequence was used as a query in wu-blastn (blast.wustl.edu, version 2.0) searches against public EST databases, and EST clones from ovary and oocytes (n = 111) were selected among the total hits displaying over 94% identity for alignments longer than 200 bp (n = 186). The ovary/oocyte-derived H3146E11-related EST clones were unambiguously mapped onto at least 11 distinct locations on the mouse genome, based on either perfect nucleotide matching or unique gene sequences

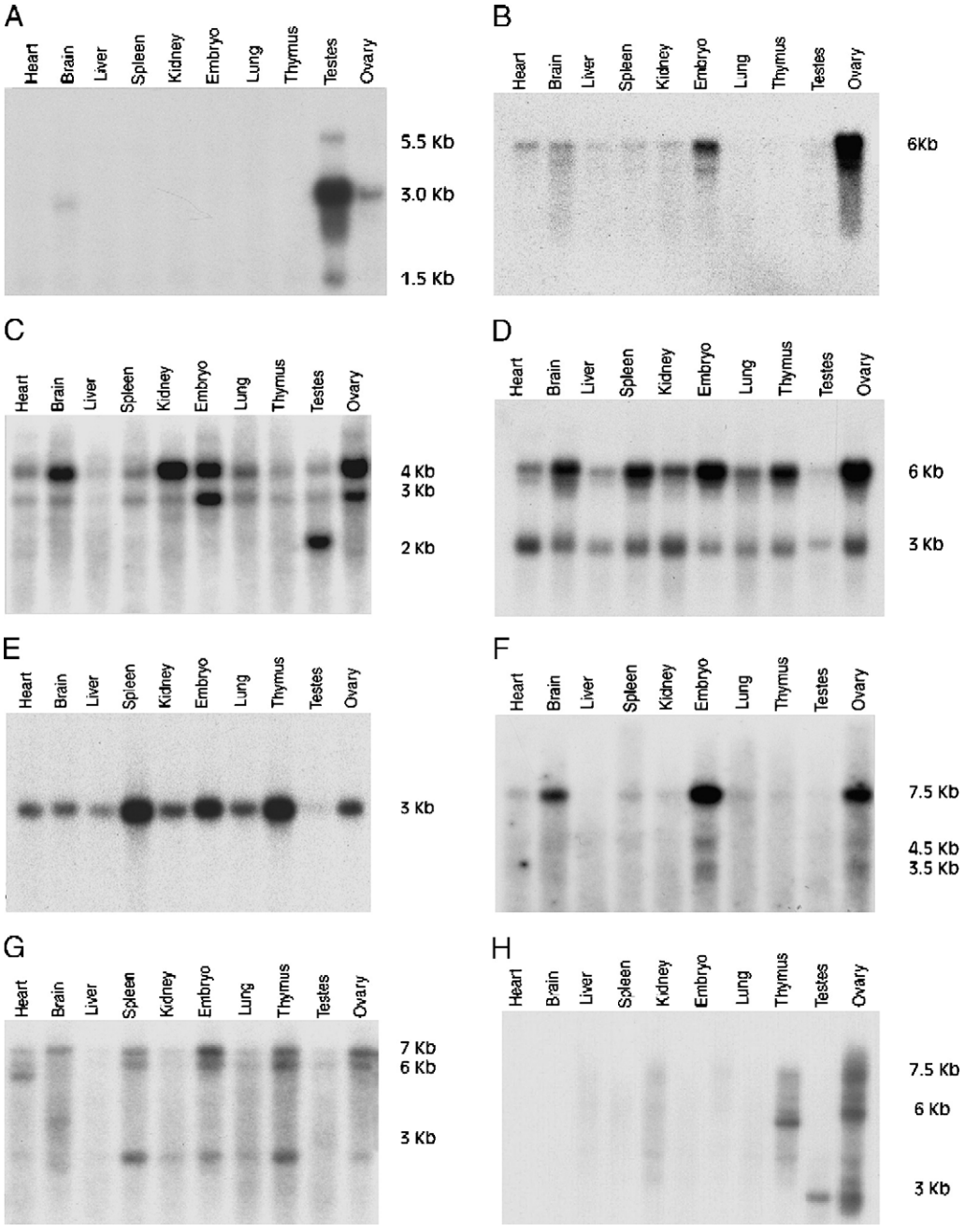

Fig. 2. Northern blots of a sample of genes detected in one or both ovarian stages by microarray. (A) *Dazl* (H3121F04); (B) the C2H2 zinc-finger *Teashirt-2* (H3083E07,H3147C07); (C) a novel CCHC zinc-finger (H3123C09); (D) a novel PHD-bromodomain (H3059E09); (E) *Dek* protooncogene (H3138A12); (F) a novel leucine rich repeat-domain (H3127G07); (G) a novel myb-domain (H3029C12); (H) a member of a novel retrosequence gene family, *RLTR10* (H3146E11).

joined by splicing events through canonical splice sites (data not shown). The total number of genomic RLTR10s was derived from the February 2002 mouse genome sequence, based on RepeatMasker annotations (genome.uscs.edu). The minimal number of nonexcised sequences belonging to this family was estimated by the minimum number of matches >90% identical over >450 nt to internal (i.e., non-LTR) portions of the H3146E11 clone obtained by wu-blastn searches of the mouse genome with overlapping 500-nt segments of H3146E11 as queries. We searched for potential autonomous RLTR10 retrotransposons by similar blastn analyses using the entire internal sequence as query; we did not detect matches containing significantly longer sequences, which would suggest a conserved pol-gag ORF.

Experimental validation

RT-PCR

One-step quantitative real-time RT-PCR with Taqman probes and primers (ABI Prism 7700 Sequence Detection

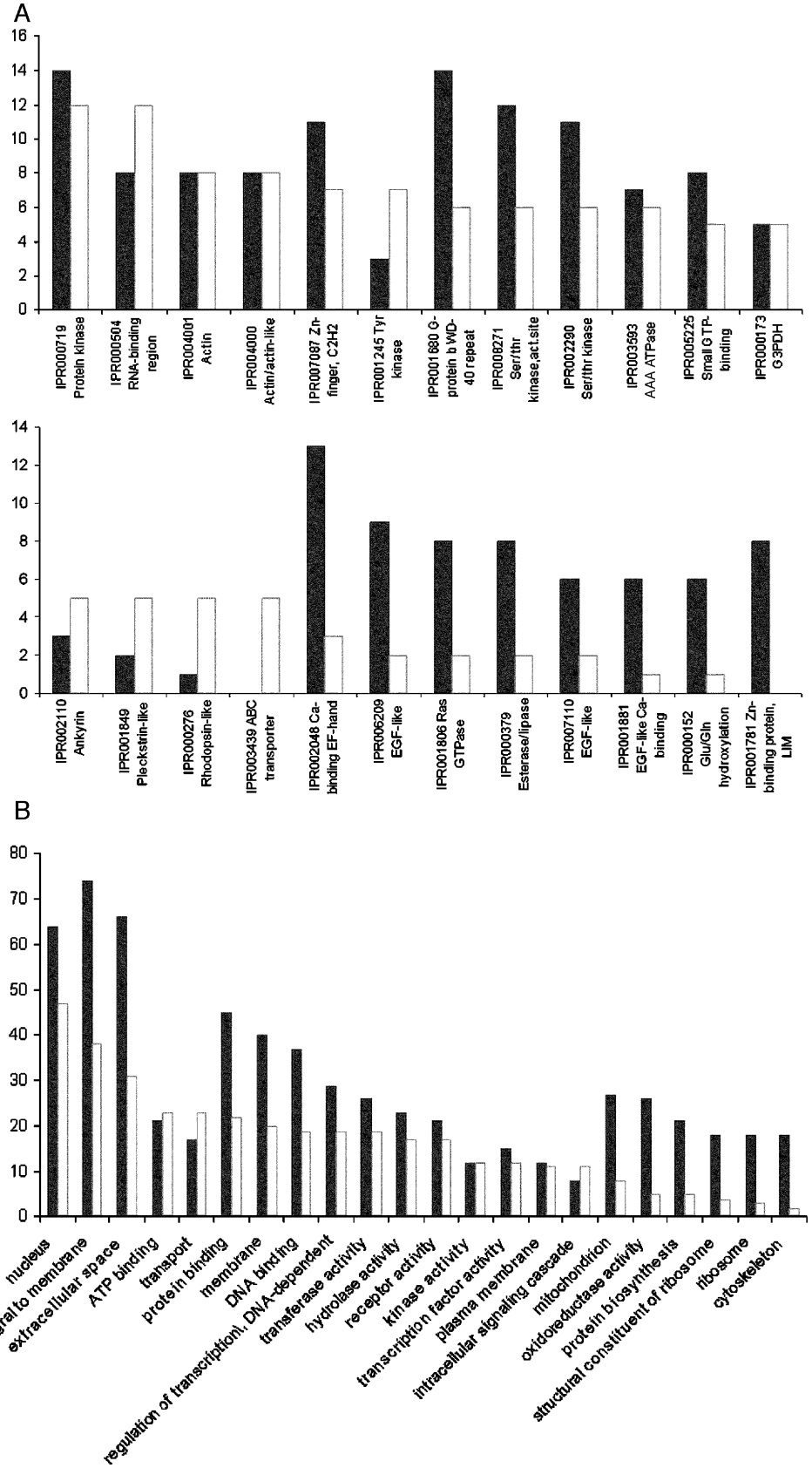

Fig. 3. Differentially expressed genes classified according to Gene Ontology terms (A) and INTERPRO domains (B) with significantly different representation between adult and newborn ovary. Light color bars, newborn ovary; dark color bars, adult ovary.

Table 4
Genes preferentially expressed in somatic cell types of newborn (A) or adult ovary (B), inferred from microarray data as in Web Tables 2 and 3 combined with published microarray results (median analysis, see Methods)

NIA clone ID	Average score	Gene symbol	Annotations
A. Inferred genes preferentially expressed in somatic cells of the newborn ovary			
H3116G07	10.4	Tm4sf6	Transmembrane 4 superfamily member 6 (Tm4sf6)
H3140D04	10.1	Lztr1	Leucine-zipper-like transcriptional regulator, 1
H3116D11	9.15	Cox7a2	Cytochrome c oxidase, subunit VIIa 2
H3140D05	8.2	Enpp5	Ectonucleotide pyrophosphatase/phosphodiesterase 5
H3130B04	7.55	Nicn1	Nicolin 1, mammalian-specific nuclear protein
H3140F05	6.35	Wfdc1	Whey acidic protein, core region Prostate stromal protein PS20
H3126F01	6.25	Plexin A2	Semaphorin/plexin repeat
H3082F02	6.1	2310014B11	Guanine nucleotide exchange domain
H3142A02	5.35	Ppp2cb	Protein phosphatase 2a, catalytic subunit, beta isoform
H3119E02	5.25	4931406N15	PMMLP homolog; IPR005844 IPR005845 Phosphoglucomutase/ phosphomannomutase alpha/beta/alpha domain I and II
H3147B01	5.25	Tyro3	Protein tyrosine kinase 3 (Tyro3)
H3140D08	4.55	Fkbp1a	FK506 binding protein 1a (12 kDa); Peptidylprolyl isomerase, FKBP-type; ABC transporter
H3124A02	4.5	Ptn	Pleiotrophin, containing PTN/MK heparin-binding domain
H3023E01	4.4	Tceb1l	Transcription elongation factor B (SIII), polypeptide 1 (15 kDa), -like (Tceb1l)
H3117D04	4.4	Fmr1	Fragile X mental retardation syndrome 1 homolog (Fmr1)
H3121A04	4.3	Ppp1cb	Serine/threonine protein phosphatase
H3121C02	4.15	Fstl1	Follistatin-like1 (Fstl)/TSC-36/Flik/FLRG; may antagonize activin
H3113F12	3.95	Prnp	Prion protein
H3122G08	3.9	2310016C16	Glutathione peroxidase
H3075F12	3.7	Tem8	Tumor endothelial marker eight precursor
H3032E06	3.65	Tulp4	Contains domains: Tubby, SOCS C-terminal, G-protein beta, and WD-40 repeat
H3031C01	3.4	Krt2–8	Keratin complex 2, basic, gene 8 (Krt2–8)
H3072A12	3.1	2810468K05	Similar subunit of nuclear pore complex NUP84
H3135C02	3	Lgals7	Galectin-7
H3002D11	2.7	C130026I10	Intron of Reticulon 4 gene
H3093E01	2.35	Rtn4	Reticulon 4 (Rtn4) (also expressed in oocytes)
B. Inferred genes preferentially expressed in somatic cells of the adult ovary			
H3131C05	8.1	4930422J18	SPla/RYanodine receptor SSB-1 homolog
H3109G09	6.9	1500036F01	Contains IQ calmodulin-binding region
H3066D10	6.65	Prkar1a	cAMP-dependent protein kinase regulator, type II (also H3145A03)
H3007E10	6.55	Ckb	Creatine kinase, brain
H3030E10	6.5	Ctsb	Cathepsin B
H3106G12	6.25	AI047808	Zn-finger, CCCH-type
H3135D11	5.95	Tpm2	Tropomyosin 2, beta (Tpm2) (other clone: H3002C07)
H3055C02	5.7	4930570C03	Uncharacterized spliced transcript
H3044D09	5.65	Scd2	Stearoyl-Coenzyme A desaturase 2 (other clone: H3025B03)
H3098A09	5.55	Epdm2-pending	Ependymin related protein-2 mRNA
H3008A09	5.5	Arhb	Ras homolog gene family, member AB
H3120G06	5.45	Tpm1	Tropomyosin 1, alpha (Tpm1)
H3008A01	5.25	Stx12	Syntaxin 12, contains SNARE domain
H3122D01	5.25	Nedd4	Neural expressed, developmentally down-regulated 4 (also in oocytes; other clone: H3104B06)
H3108C10	5.1	Mm.260347	Composite SINE repeats
H3056C12	5.05	Dst	Dystonin; hemidesmosomal plaque protein
H3132G02	4.95	App/Mm.131766	Amyloid beta (A4) precursor protein
H3054E05	4.9	Rps27	Ribosomal protein
H3076C06	4.9	Mm.182249	Composite SINE repeat
H3134F06	4.9	Sparcl1	SPARC-like 1 (mast9, hevin)
H3113G03	4.8	Agrn	Agrin
H3055D08	4.75	Cop1	Constitutive photomorphogenic protein, contain ring-type zinc finger
H3078H12	4.5	Sulf2-pending	Sulfatase weakly similar to N-acetylglucosamine-6-sulfatase
H3106B04	4.45	Mm.27490	Transmembrane protein similar to CGI-40 protein
H3148G03	4.4	Dkk3	Dickkopf homolog 3; Wnt pathway
H3139A10	4.2	Ckap1	Tubulin-specific chaperone b/tubulin folding cofactor b; contains CAP-Gly domain

(continued on next page)

Table 4 (*continued*)

NIA clone ID	Average score	Gene symbol	Annotations
B. Inferred genes preferentially expressed in somatic cells of the adult ovary			
H3146E06	3.45	Mcam	melanoma cell adhesion molecule (Mcam)
H3112A08	3.4	Matn2	Matrilin 2
H3115C09	3.35	Lpp	Cytochrome c heme-binding site; Zn-binding protein, LIM
H3133C01	3.35	Sui1-rs1	Suppressor of initiator codon mutations, related 1
H3151G01	3.35	Sf3b1	Splicing factor 3b, subunit 1, 155 kDa
H3154A09	3.25	Fln1	Filamin a/filamin 1 endothelial actin binding protein abp x280 nonmuscle
H3155A05	3.2	Tuba2	Tubulin alpha 2 MGD|MGI:107804
H3028G03	3.15	Robo1	Roundabout homolog 1 (Drosophila)
H3148D05	2.95	Dcn	Decorin, contains Cysteine-rich region and leucine-rich repeat
H3154H04	2.75	Aplp2	Amyloid beta (A4) precursor-like protein 2 (Aplp2)
H3153C09	2.55	Timp2	Tissue inhibitor of metalloproteinase 2
H3115G02	2.15	Cadh	Cadherin

The average scores were obtained with Focus for the newborn to adult comparison (see Methods and WEB Tables 2 and 3).

System, Applied Biosystems, Foster, CA, USA) was performed to confirm microarray results. PCR was for 40 cycles according to the manufacturer's protocols. Reactions were normalized to Gapdh (except RLTR10, H3146E11, which was normalized with β-actin). Each set of reactions included RT-minus or nontemplate samples to control for nonspecific amplification. The forward primer, reverse primer, and probe sequences were, for Tsh3, 5′-TCGCGGGAGAAGCATCATG, 3′-CGCTTGTTCCTCCGGCTCTA, and GGCAGCTGCCTTTAACTCATC; for *Tsh2*, 5′-TCGTGTTGGAGGCTGTAGAA, 3′-TCACCCGAACACTCTCA, and 5′-CAGAACTATTC-TTCATCCACGTCTGCTA; for *Dazl*, 5′-GTAGAAGTCTGTGGACCGAA, 3′-TTAAGCACTGCCCGACTTCTTCTG, and 5′-CAGACAAGAGACCACTGTCTGTATG; for *Wnt4*, 5′-TGGGAAGGTGGTGACACAAG, 3′-ATGCCCTTGTCACTGCAAAG, and TGTATACGCCATCTCTTCAGC; for *Mllt10* (H3076A09), 5′-GCAGATCCCCATCAACAATC, 3′-GCTACAGGCCCAGTTTTATC, and 5′-TCCACTCCACACAGCTACCACCAA; for *Fstl1* (H3121C02), 5′-GCAGAATGAAACAGCCATCA, 3′-AGGCACTTGAGGAACTCTTG, and 5′-CCTCTGTGTTGACGCCCTCATTG; for *Sf1* (H3053B11), 5′-CAAGAGTTAGTGCTCCAGTTG, 3′-ACGAGGCTGTGGTTGTTCAG, and CGTCTGTCTCAAGTTCCTCATC; for *Pole4* (H3-126A12), 5′-ATCGCAAAAGATGCCTACTG, 3′-CAG-ACCATCCAGCTCCACAA, and 3′-CGAATTCATCCACAGCTTCTATTGCA; for *Taf7l* (H3-154D11), 5′-GGAAGCTTAACATAATGGAA, 3′-TCATCTGGAGGGTGAATCCA, and 5′-AAAGCCTCAGCCTGGCACACCAA; for *Podocalyxin* (H3029C03), 5′-CTACTGTCGCCTGCATCTCA, 3′-GTGACCAGTTGTGACTTCTG, and 5′-TGTCCAGAGCCACCAAAGTGCCAC; for *RTLR10* (H3146E11), 5′-TGATAACTCCCTGGGCATGT, 3′-CCCGTCCTTGGGTGACAA, and 5′-CCAACCTAAGACAGGGATCAAACCAATGC; for *Piwil2* (H3149H10), assay Mm005002383 (TaqMan® Gene expression assays, Applied Biosystems); for *Nxf2* (H3145H01), 5′-CCTCACGTTAAGATCCTGAA, 3′-CCTTGCCATTCTGAAGATTTC, and 5′-CCTTCTGCAACTGCTTCTTGGA; for *Tktl1*, 5′-ATTATCACCGTGGAGGATCA, 3′-ACTACGAGGCACATCCATTA, and 5′-AGCAGTCTCCATGGAGCCTAAC; for β-actin, ACGGCCAGGTCATCACTATTG, CAAGAAGGAAGGCTGGAAAAGA, and CAACGAGCGGTTCCGATGCCC; and for *Gapdh*, Rodent GAPDH control reagents (VIC Probe) (Applied Biosystems). Real-time PCR was performed with serially diluted RNA (1.5, 1.0, 0.50, 0.25, and 0.1 ng as template). Reaction mixtures were at 95°C for 3 min, followed by 40 cycles at 95°C for 30 s, 62°C for 45 s, and 72°C for 1 min.

In situ hybridization

In situ hybridization was performed on paraformaldehyde-fixed samples either embedded in paraffin or frozen as described previously (Wilkinson and Nieto, 1993). To linearize the pSPORT 1 plasmid containing each cDNA insert, 15 μg plasmid DNA was digested overnight with *Sal*I or *Not*I. DNA was purified by three phenol/chloroform extractions followed by chloroform extraction and ethanol precipitation. Sense- and antisense-labeled transcripts were generated using 1 μg of DNA and RNA polymerases (T7 and SP6) according to the manufacturer's protocols (Roche). Signals were scored after 5–6 h in nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate. *RLTR10*, *Gdf9*, *Deltex2*, *BicD2*, and *Zfp37* probes were derived from constructs of the NIA cDNA library (Tanaka et al., 2000; GenBank accession numbers BG087929, BG080778, BG076683, BG080326, and BG080774).

Immunohistochemistry

Samples were fixed in either 4% paraformaldehyde-PBS or Histochoice (Amresco, Ohio) and embedded in paraffin. Sections were unmasked in citrate buffer and stained using Dako LSAB2 system (peroxidase and DAB) or Alexa-dye fluorescent secondary antibodies (Molecular Probes). Primary antibodies were as follows: mouse antihuman cyclin D3 (Stressgen), goat anti-mouse c-kit (Santa Cruz), mouse antihuman cytokeratin 18 (Chemicon), rabbit antihuman PDGF receptor β (Upstate), and rat anti-mouse laminin A1 (Chemicon). Images in Figs. 5A–D and 6 were taken with

an Axiovert 200 microscope (Zeiss, Gottingen, Germany) coupled to a Spot® camera (Diagnostic Instruments, Sterling Heights, MI). Images in Figs. 5E–F were taken with a HiRes CCD camera and processed on a Deltavision system v.5.10. All images were processed with Adobe Photoshop® v.8.0.

Northern blot analyses

FirstChoice™ Northern blot mouse blot I (Ambion, Inc., Austin, TX) multiple tissue membranes were used with probes labeled with the RediPrime™ II labeling System (Amersham Pharmacia) according to the manufacturer's protocol; comparable amounts of RNA in each lane were confirmed with a β-actin control. About 100 ng of cDNA was the template for each probe synthesis. The activity of α-^{32}P in each hybridization was 10^6 cpm. Templates for probe synthesis were cDNA inserts from NIA mouse cDNA libraries. All cDNA inserts were excised from plasmids with *Not*I and *Sal*I and reamplified by PCR before use. The primers used were, for Dazl1 (H3121F04), 5′-GGGCTTCGGTCAAATTTGCTATG and 3′-GGTGGTGATAGGCTGGGAATG (250 bp); for C2H2 zinc-finger Teashirt-2 (H3083E07,H3147C07), 5′-CGAAGAGGACACAGACTCTA and 3′-GAGGTCATCACTGAGTCCAA (135 bp); for novel CCHC zinc-finger (H3123C09), 5′-AGCGGAGGAGGAGTCACA and 3′-CACTTTCGTCTCCCTGGAAA (503 bp); for novel PHD-bromodomain (H3059E09), 5′-GAGCAGGACCCCGAGAAGAA and 3′-CGTTTGGGCTGCCAGACAAG (136 bp); for Dek proto-oncogene (H3138A12), 5′-CTGACGATGCAAGTGTCTTC and 3′-CTACTAGACTCATCTGACAG (470 bp); for novel leucine rich repeat domain (H3127G07), 5′-GGACTTCTCTTAACAGTAC and 3′-CAGGCCAGGAGTTTTGACAC (390 bp); for novel myb domain (H3029C12), 5′-AGTCAGGAGGGTGAGTCTGA and 3′-TGTGCATTTGGGTGCTTGAC (430 bp); for member of a novel retrotransposon gene family, RLTR10 (H3146E11), 5'-TGATCAATAGATCCTCGCTGTG and 3'-TCGATCCTCTCCAAATATCCAG (364 bp).

Results

Based on the normalization and selection criteria outlined in Methods, a total of 9582 transcripts showed significant expression in at least one of the two stages of ovarian development tested (WEB Table 1). Northern blot hybridization confirmed expression in adult ovary of a sample of eight genes that were either known to be expressed at low levels (*Dazl*, Fig. 2A) or were not previously reported in the ovary (*Tsh2* and *Dek*, Figs. 2B and E) or were novel/uncharacterized (Figs. 2C, D, and F–H).

Using a sensitive statistical test (see Methods, Fig. 1 and WEB Tables 2 and 3), we identified 2247 transcripts (2115 UniGene clusters) that were differentially expressed in newborn compared to adult ovary. Mapping the full set of protein-encoding genes on the Gene Ontology database (see Methods and Fig. 3A) suggested extensive functional differences. They largely reflected a vast repertoire of genes involved in regulating nuclear activity and cell–cell interactions during morphogenesis in the newborn ovary and a greater prevalence of growth-related processes in the adult. Thus, genes associated with transcription regulation, ATP binding, and intracellular signaling functions were enriched in expression in the newborn ovary (Fig. 3A). The distribution of INTERPRO protein domains (Fig. 3B) showed an especially high number of clones encoding RNA binding factors. In contrast, in the adult (Fig. 3A), increased expression was seen for genes involved in growth-related processes, including oxidoreductase activity and protein biosynthesis (Fig. 3B). Table 2 lists a set of genes with known or potential developmental roles that show differences in expression levels between newborn and adult ovary.

The analysis then proceeded with tests that validated the microarray results and integrated them with complementary data available in public databases, providing a list of genes with different stage and cell specificity.

Validation

Standard and real-time RT-PCR results confirmed that results were reliable for the preponderance of genes detected. For example, of 12 genes tested, the relative levels for three known to be expressed in ovary (*Dazl, Wnt4, Sf1*) and seven that are novel or previously uncharacterized in this organ (*RLTR10*/H3146E12, *Piwil2*/H3149H10, *Tktl1*/H3045F08, *Taf7L*/H3154D11, *Mllt10*/H3076A09, *Podx*/H3029C03, and *Fstl1*/H3121C02) were confirmed (*t* test, $P < 0.05$; Fig. 4 and data not shown). Transcripts assigned to only 2 of the 12 genes showed mean intensity values that were not significantly different by real-time RT-PCR (*Pole4*/H3126A12 and *Tsh2*/H3083E07; *t* test, $P > 0.05$). Further investigation of one of them, *Tsh2* (Caubit et al., 2000), indicates that the discrepant values likely reflect further complexity in the results. The mean intensity difference for *Tsh2* was opposite that expected from microarray data; but this could result from cross-hybridization with a paralogous gene (*Tsh3*, >200 nt with 79% identity over the cDNA clone H3083E07). Consistent with this possibility, *Tsh3* expression in the ovary showed a mean intensity fold ratio reversed compared to *Tsh2* (WEB Fig. 1).

An independent computer-assisted analysis of redundant clones present in the 15K microarray showed a comparable result: under the stringent assumption that all UniGene clusters contained truly coexpressed sequences (i.e., no differentially expressed isoforms or inclusion of more than one gene in an entry), 80% of the genes detected were concordantly assigned among the clones detected.

Additional transcripts were studied by in situ hybridization or immunohistochemistry, including four controls

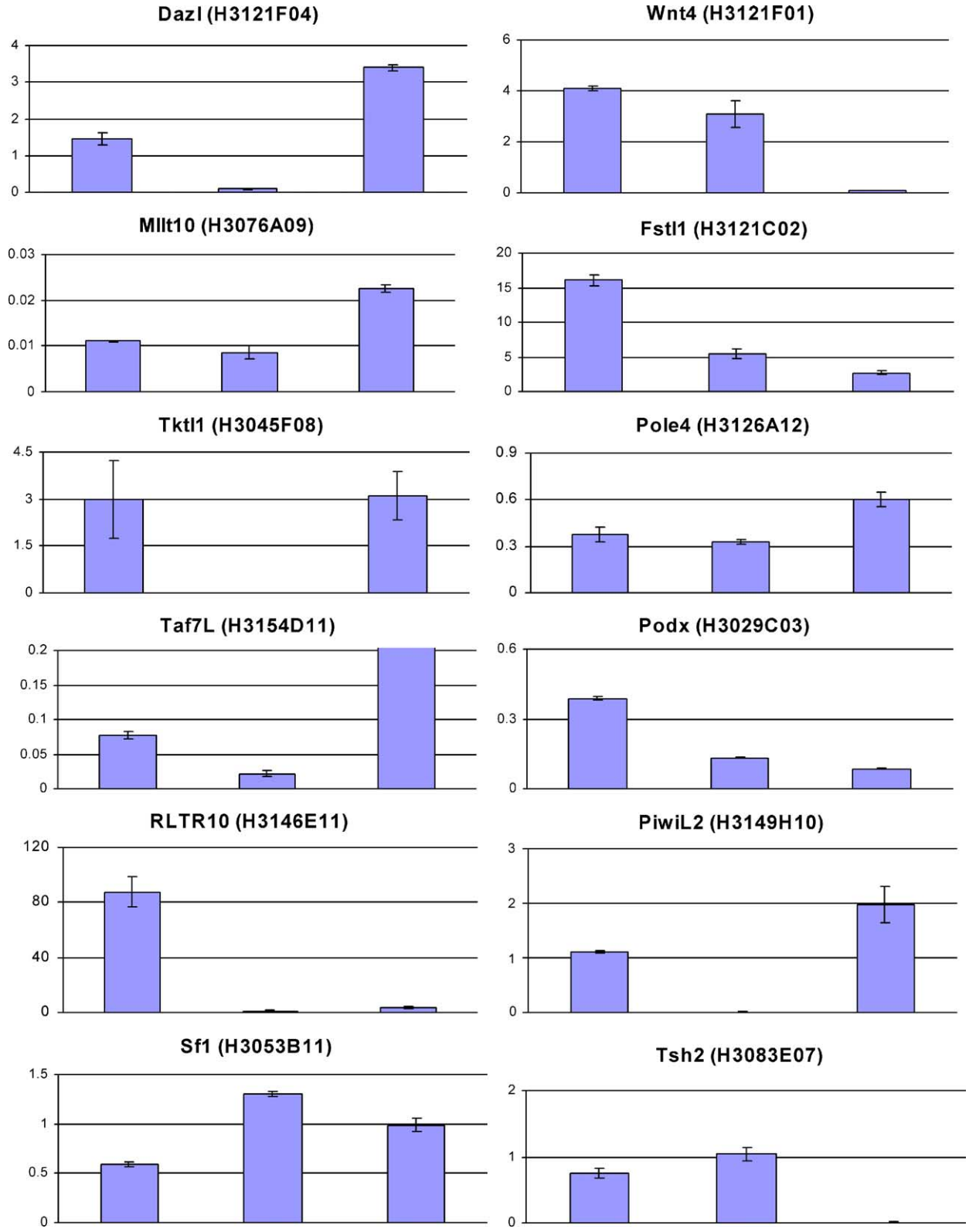

Fig. 4. Histograms representing relative mean expression values and standard deviation for 12 probes by real-time PCRs (y-axis). From left to right: newborn ovary, adult ovary (8 weeks) and adult testis. Values were normalized to GAPDH mean expression.

(for *cyclinD3, c-kit*, *Gdf9*, *Bmp15*) and four others that were previously uncharacterized or unreported in the ovary (*RLTR10*, *Dtx2*, *BicD2*, *Zfp37*, Fig. 5 and data not shown). These data provide independent experimental validation of the microarray analysis and are discussed in greater detail below.

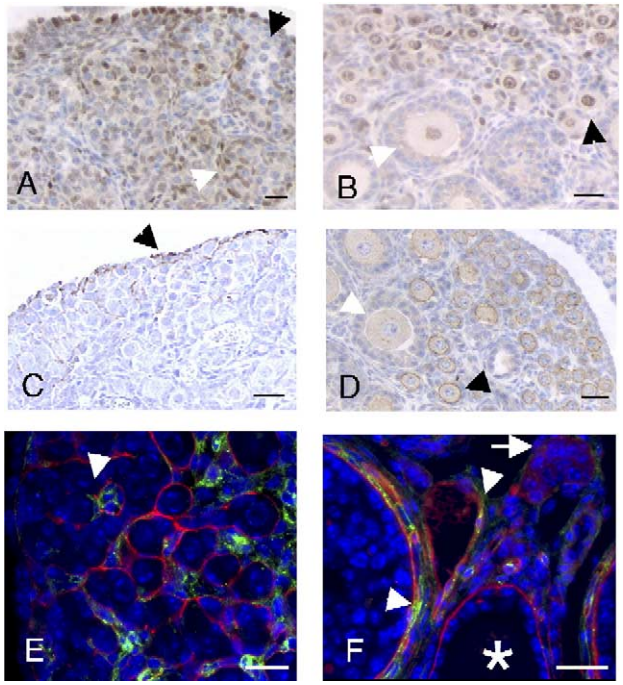

Fig. 5. Immunohistochemistry of newborn ovary enriched genes. CyclinD3 is expressed in somatic cells (A, white arrowhead) but not oocytes (A, black arrowhead) before follicle formation, then it extends to oocytes as primordial follicles form (B, black arrowhead) and is progressively lost in growing follicles (B, white arrowhead) of the maturing ovary. CK18 is expressed in ovarian surface epithelium (black arrowhead) and pregranulosa cells that infiltrate the oocyte clusters in the newborn ovary (C). C-kit is expressed preferentially at the membrane of primordial and early primary follicles (D, black arrowhead) and lost at the secondary follicle stage (white arrowhead). Counterstaining for A-D: hematoxylin. PDGFR-β is detected in stroma and vessels of newborn ovary (E, green), partly coincident with follicle formation (white arrowhead shows fragmentary basal lamina detected by anti-laminin A1 antibody, red). In the maturing ovary, it persists selectively in theca interna of healthy follicles and some vessels (white arrowheads), but not interstitial glands (white arrow) or atretic follicles (asterisk). Scale bars: 20 μm.

Stage-specific processes are reflected in the microarray

Even in the absence of information assigning individual genes to particular cell types, whole organ expression profiling can help to investigate complex time-related differences that are difficult to model in vitro. In addition, the approach is directly justified for genes that are expressed either in several cell types simultaneously or in complex sequential patterns, but preferentially at a particular stage. For instance, cyclin D3 (H3041D11), which was correctly assigned as enriched in newborn ovary by the microarray analysis, is expressed in fetal and newborn pregranulosa cells, then appears in oocytes at the time of primordial follicle formation, and is then down-regulated in both oocytes and granulosa cells during follicle growth (Zhang et al., 1999, and confirmatory Figs. 5A–B).

Other markers of early ovarian differentiation include cytokeratins (CK18, CK19, and CK8, Appert et al., 1998), the clones for which were all correctly found enriched in the newborn ovary (H3022G09, H3021B02, H3007G06, H3031C01, H3036G12, H3104F03, and confirmatory data for CK18 in Fig. 5C). These data suggest that some of the uncharacterized transcripts identified in our study may similarly show stage-specific expression in the ovary. For other genes, however, whole organ gene expression differences rather reflect relative changes in cell-type composition. We addressed this possibility as follows.

Cell type-specific expression patterns

Cell type-specific expression data are available for adult ovary and usually not for newborn ovary, largely for technical reasons. However, we combined information from NIA EST libraries and publicly available microarray data to generate a partial list of genes enriched in oocytes (Tables 3A–B) and in somatic cells (Tables 4A–B) (see Methods and Results below). We were thus able to compare changes in gene expression levels with variations in cell type numbers that accompany ovary development (Sforza et al., 1993). In particular, as the ovary matures, the volumetric fraction of oocytes and granulosa cells from small follicles goes down, while stroma, steroidogenic cells (of both stromal and granulosa origin), and vessels go up. In rodents, these variations are particularly marked, because primordial follicle formation occurs largely within a restricted window of time around postnatal day 2 (e.g., Peters, 1969; Rajah et al., 1992). Ovarian histogenesis and maturation are thus discretely phased. Consequently, oocyte-specific and small follicle-specific genes are expected to be considerably enriched in the newborn ovary, and steroidogenic and endothelial genes in the adult. In particular, a gene enriched in adult (ovulated) oocytes compared to adult soma and enriched in newborn compared to adult ovary is likely oocyte specific throughout ovarian development (Table 3A). Conversely, a gene enriched in ovary compared to oocytes (see Methods), as well as in the adult ovary compared to the newborn, is likely associated with somatic cells involved in follicle growth (Table 4B).

Microarray data on stage specificity are particularly informative when variations in gene expression patterns contrast with those expected from the associated cell type, because such contrasting patterns are directly suggestive of a stage-specific role. For instance, genes specifically expressed in oocytes and involved in oocyte growth were found to be enriched in the adult, in spite of the relatively small number of oocytes compared to organ size at this stage (e.g., *Gdf9* (H3056F01), *Bmp15* (H3056A05), and *Zp3* (H3102C08); and novel genes in Table 3B, and Figs. 5C and E–F). Conversely, when they are enriched in newborn ovary, genes from steroidogenic, stroma, and vascular cells are candidates for a role in primordial follicle formation or maintenance. In this case, the few genes reported with such an expression pattern in ovary (e.g., *NGFR*, Dissen et al., 1995) were not available on the microarray. However, we have identified genes belonging to

this class that were previously uncharacterized in the ovary (*Podocalyxin*, *Agtr2*, *PDGFRβ*, see below). Inferred markers of early differentiation of somatic lineages are given in Table 4A. These genes were studied further, as follows.

Cell type-specific transcripts enriched in newborn ovary: known genes

Two genes of central importance for ovarian development were enriched in the newborn ovary: *c-kit* (H3136A01, consistent with a reported 2-fold decrease in mRNA expression in growing oocytes, Manova et al., 1990; and Fig. 5D) and *Wnt4* (H3121F01) (confirmatory data in Fig. 4). These genes are expressed in different cell types (predominantly in oocytes or granulosa cells, respectively), but are both associated with severe phenotypes affecting perinatal ovarian differentiation (*c-kit*, Matsui et al., 1990; *Wnt4*, Vainio et al., 1999).

Germ cell-specific genes enriched in newborn ovary show more uniform expression in oocytes throughout ovary maturation, including *Dazl* (H3121F04, H3132H08, Fig. 4), expressed in oocytes at all follicle stages and required for progression through meiotic prophase in both sexes in the mouse (Ruggiu et al., 1997); *Nobox* (H3157G03), required for early oocyte growth and oocyte maintenance (Rajkovic et al., 2004); *Mater* (H3102E07, Tong and Nelson, 2000), a maternal effect gene expressed from oocytes at all follicle stages; and *Zfr* (H3076B02), encoding a zinc finger RNA binding protein with highest levels in prophase I germ cells (Meagher et al., 1999). Taken together, these data suggest that genes identified as newborn ovary enriched by the microarray analysis are candidates for functions in primordial follicle formation and/or early follicle differentiation.

Cell type-specific transcripts enriched in newborn ovary: novel protein-encoding genes

The list of oocyte-enriched genes included genes not previously reported in the ovary but with known roles highly suggestive of enrichment in oocytes. One example is the checkpoint kinase 1 homolog (H3092A01) that is highly expressed in meiosis I mouse spermatocytes. *Chek1* encodes a protein that colocalizes with *Atm* (OMIM607585), associated with arrest in meiotic prophase in both sexes when mutated (Flaggs et al., 1997). Similarly, the LAP/PHD finger *Mllt10/AF10* (H3076A09) and the TATA-binding protein associated factor *Taf7L* (H3154D11) (both confirmed by real-time RT-PCR as stronger in newborn ovary, Fig. 4) were reported as nearly testis specific (Linder et al., 1998; Pointud et al., 2003), but are more likely present in all germ cells.

The list of somatic genes enriched in the newborn (Table 4A) included a glycoprotein containing a follistatin-like domain that can imitate follistatin effects in embryonic axis induction but had not been reported in the ovary (*Flst1/Flik/TSC-36*, Patel et al., 1996; Shibanuma et al., 1993, confirmed by quantitative RT-PCR, Fig. 4). Other genes were known to be expressed in the ovary, but their developmental expression pattern is still poorly characterized. They included the fragile X syndrome gene *Fmr1*, associated with premature ovarian failure in humans and reported to be highly expressed in granulosa cells (Hergersberg et al., 1995). In addition, the enrichment in the newborn of a lectin known to be specific to ovary and adult-stratified epithelia (*Galectin-7*, Magnaldo et al., 1998, confirmed by RT-PCR in data not shown) may be of interest with regard to the epithelial origin of pregranulosa cells (e.g., Rajah et al., 1992).

Vascular markers were identified on the basis of gene annotations and lists of genes available in expression profiling reports (Ho et al., 2003; Huminiecki and Bicknell, 2000). Two genes previously uncharacterized in the ovary were found to be enriched in newborn. One is highly specific for early events of developmental vasculogenesis and is later expressed in podocytes (*podocalyxin*, H3029C03; Hara et al., 1999; confirmed by real-time PCR, Fig. 4); another is upregulated during fetal angiogenesis with restricted expression pattern in adults (*Agtr2*, or angiotensin II receptor type 2, H3125F10, de Gooyer et al., 2004; Kakuchi et al., 1995). Both are associated with urinary tract anomalies leading to neonatal lethality when inactivated (Doyonnas et al., 2001; Nishimura et al., 1999), and they may have a comparably important function in ovarian development. In a similar way, traditional stromal markers in newborn ovary included a gene (H3140C02, *PDGF receptor-β*) that is also involved with urogenital morphogenesis (Puglianiello et al., 2004). We verified that PDGFR-β is detected in most stroma and some endothelial cells of the newborn ovary, around newly formed follicles, primary follicles, and in a few cells that intercalate oocyte clusters, often coincident with laminin deposition (by double staining with specific antibodies, Fig. 5E). In the maturing ovary (Fig. 5F), it is still high in some vessels, in the residual cortical stroma, and then confined to theca interna cells of some healthy follicles up to the antral stage. In contrast, most of the stroma and all nonstromal cells are negative or very faintly positive. This is consistent with upregulation in the newborn and suggests an early role for *PDGFR-β* in the differentiation of ovarian stroma.

Cell type-specific transcripts enriched in newborn ovary: untranslated transcripts expressed in oocytes of primordial and primary follicles

Among the transcripts not previously known to be expressed in oocytes and/or granulosa cells of small follicles, we identified a large number of retroviral-like sequences and other repeats (36% [814/2247] of the clones contained sequence repeats; WEB Tables 2 and 3). Several of them were differentially expressed between newborn and

adult ovary, reminiscent of reports identifying differential expression of this class within the ovary and between ovary and testis (see Discussion). They included a large number of clones representing a novel retrosequence family (RLTR10) (H3146E11 and others) that was 20- to 100-fold more expressed in newborn ovary (Figs. 3–5 and data not shown). In situ hybridization of the probe to sections of the ovary showed strong expression in germ cells of the nascent primordial and primary follicles in the newborn (Fig. 6A). In corresponding sections of adult ovary, a gradient in expression can be seen in follicles at various stages of maturation (Fig. 6B). In the same series, diffuse expression in granulosa cells lying nearest the oocyte may represent "leakage" from the oocyte (Canipari, 2000).

The cDNA clone is essentially identical to a ~2.8 kb genomic retrosequence on mouse chromosome 12 flanked by long terminal repeats (LTRs) of the RLTR10 class (www.girinst.org/Repbase_Reports.html). The internal sequence displays three short segments of significant similarity to classical IAP retrosequences (e.g., in accession numbers M17551 and M73818), in a novel murine repeat family (schematically aligned with several cloned variants in WEB Fig. 2). No significantly similar sequence is detected in the human genome. H3146E11 has the size and sequence hallmarks of nonautonomous, internally truncated retrosequences that are thought to be occasionally generated from canonical ~6kb IAP-like elements (reviewed by Kuff and Lueders, 1988). We failed to detect full-length IAP sequences having high similarity to H3146E11, but about 3800 sequences contain this class of LTRs. In addition, we observed many sequences similar to internal portions of H3146E11; about 500 loci show >95% identity over at least 250 nt and ovary/oocyte-derived ESTs can be assigned to at least 11 distinct RLTR10 loci (see Methods). Therefore, the H3146E11 transcript proved to represent a whole family of sequences, all relatively specific for germ cells of nascent follicles, with some expression observed in some granulosa cells of mature follicles.

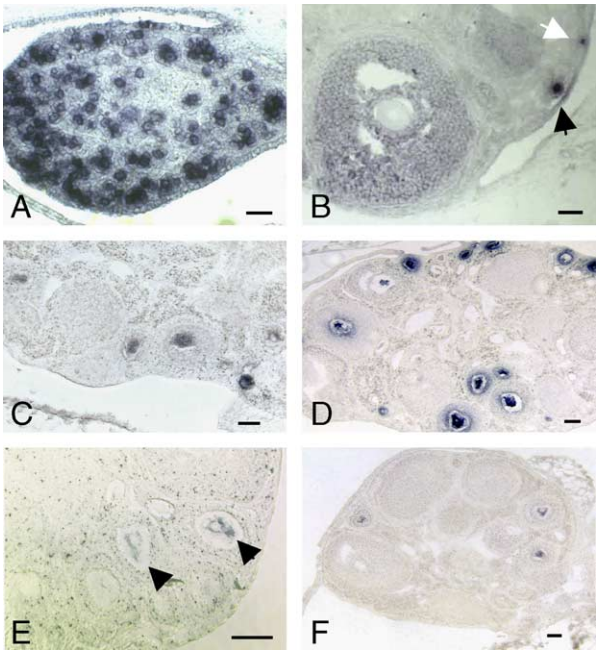

Fig. 6. mRNA in situ hybridization for genes enriched in the newborn ovary (A and B) and in the adult (C–F). The *RLTR10* retrosequences show strong signals in oocytes of nascent follicles in the newborn ovary (A). The signals persist in primordial (white arrow) and primary follicles (black arrow) in the adult, but are lost in growing follicles (B). The transcripts of *Deltex* (C), *Gdf9* (D), *BicD2* (E), and *Zfp37* (F) are detected only in growing oocytes. Scale bars: 40 μm.

Cell type-specific transcripts enriched in adult ovary: known genes

In maturing ovary, we did not attempt to distinguish specific regulators of follicle maturation from genes essentially associated with expansion of adult cell types, although data on known oocyte-specific genes were consistent with our prediction (for example, *Gdf9* and *Bmp15*; Fig. 5H and data not shown). Markers of ovarian follicle growth were not confined to oocytes but were also enriched in the somatic compartment. This would be expected from the estimated 21 doublings of granulosa cells during follicle growth, with the basal lamina at the interface with stromal theca cells increasing 2^{19}-fold (Rodgers et al., 2001). More abundant expression of general basal lamina components in the adult ovary thus included *Col4a1* (H3159G06) and *Col4a2* (H3138E02).

Markers for growing follicle cells also showed higher levels in adult ovary (e.g., *InhibinA* (H3142C10), Drummond et al., 1996, and *Foxo1* (H3065B05 and H3100B07), Shi and LaPolt, 2003). In addition, the adult ovary gene set included all relevant steroidogenic regulators and enzymes that are expressed in luteal cells and adult stroma (including steroidogenic factor *SF1* (H3053B11, confirmatory data in Fig. 4), which regulates transcription of all cytochrome *P*450s; P450ssc/Cyp11A1 (H3094D01), catalyzing the conversion of cholesterol to pregnenolone; steroidogenic acute regulatory protein StAR (H3097G04), which regulates accessibility of cholesterol to P450ssc; P450c17/Cyp17 (H3018A11), a theca cell marker that catalyzes the formation of dehydroepiandrosterone (DHEA) and is upregulated by Gdf9 in preantral follicles (Vitt et al., 2000); and two clones (H3059B01, H3070C08) for placental 3-β-hydroxysteroid dehydrogenase Hsd3b1 (see OMIM109715)). Interestingly, consistent with its role as a repressor of P450c17/Cyp17 (Shibata et al., 2003), the C4-type zinc-finger Coup-TF2 was upregulated at birth (H3124H07). In addition, adult ovary showed an enrichment of ubiquitous vascular markers, consistent with massive cyclical angiogenesis as a feature of the mature ovary. They included *Selectin* [Selel/Mg160, H3052F04], *C1qr1* [H3146C06], *Pecam* [H3052F06], and *Eg1* [H3017G10] Liu et al., 2002 Yamaguchi et al., 2003). Finally, novel genes expressed in oocytes showed expres-

sion patterns indicative of a role in late stages of oogenesis, as follows.

Cell type-specific transcripts enriched in adult ovary: novel genes

In spite of the small volume of oocytes in the adult ovary, microarray analysis detected transcripts expressed only in growing oocytes. The combined usage of microarray information with NIA EST library data (see Methods) detected both well-known genes upregulated in growing follicle oocytes (*Gdf9* and *Bmp15*) and also additional genes previously uncharacterized (Table 3B). In situ hybridization verified the results for three that were not previously characterized. One, *BicaudalD2* (H3054C10; low expression by ISH in Fig. 6, consistent with EST counts, Table 3B), a cytoplasmic-helical coiled-coil protein involved in microtubule polarization, is orthologous to a *Drosophila* gene essential for the establishment of oocyte identity and embryonic axis polarization (Hoogenraad et al., 2001 and references therein). A second, *Zfp37* (H3056E09; confirmed by ISH in Fig. 6), a C2H2 KRAB zinc finger gene, is also highly expressed in the adult testis, with highest abundance during late spermatogenesis and spermiogenesis (Burke and Wolgemuth, 1992; Mazarakis et al., 1996). The third example, *Deltex2* (H3004E12; confirmed by ISH in Fig. 6), is a cytoplasmic SH3-ringH2 protein that mediates Notch receptor signaling (Kishi et al., 2001).

Finally, microarray data suggest that additional genes that are downregulated in female germ cells compared to ovary and uterus (inferred by comparison to published data, as above) are likely expressed in growing follicle cells (Table 4B).

Discussion

The overall census of about 9000 cDNAs significantly active in the ovary extends earlier results with cultured follicles or granulosa cells, which included the identification of verified or inferred targets of critical ovarian determinants (Burns et al., 2001; Dong et al., 1996; Rajkovic and Matzuk, 2002; Vanani et al., 2002) and two pilot microarray analyses that showed 15 of 588 known genes changing in level during late follicle maturation (Liu et al., 2001) and a comparable number of changes in follicles from women with full vs. depleted complements (Chin et al., 2002).

Selective expression in female gonads has, however, been inferred for only a few genes (for example, the homeobox gene *Emx2*, OMIM 600035, Pellegrini et al., 1997; and nuclear steroid hormone receptor-related factor *Dax1*, OMIM 300200, 300018). Pilot comparative screenings for embryonic gonadal genes have been carried out based on ESTs (Takasaki et al., 2000) or clone arrays (Grimmond et al., 2000); by high-throughput in situ hybridization (Wertz and Herrmann, 2000); or by cDNA difference sampling (Perera et al., 2001). But those studies detected mainly a number of testis-specific genes, whereas genes for early stages of ovarian development were poorly represented. In a review, Koopmann et al. (2002) even questioned whether organ-specific transcriptional activity in the ovary could be systematically detectable by microarrays; but our findings show that time- and cell type-dependent markers and pathways are definable for ovary as in other developmental models.

Using whole organ preparations for microarray analysis provides information on cellular interactions and processes—in this case, follicle formation and maturation with regulated control of quiescence and growth—for which no in vitro model is currently available. In addition, taking into consideration variations in cell populations that take place with ovary maturation, this design permitted us to confirm cell type-specificity in newborn ovary or to suggest stage specificity for cell type-specific genes in newborn or adult (see Results). Finally, the restricted time interval of follicle development in the murine ovary enabled us to sort out genes with selectively high expression in primordial and early growing follicles (high in newborn ovary) from genes enriched in growing follicles (high in adult ovary). Although the microarray analyses cannot readily distinguish between changes in cell population sizes and changes in expression levels within a differentiating cell lineage (see above), further experiments and independent sources of evidence permitted the discrimination of genes that were expressed throughout oogenesis from others selectively involved in primordial follicle development or maintenance (see Results).

cDNAs that showed differential expression as a function of development or follicle differentiation were of particular interest. EST correlations provided a list of oocyte-specific genes with differential expression at different stages of folliculogenesis. Similarly, a whole organ approach providing a useful entry to characterizing primordial follicle formation, a notoriously complex morphogenetic process involving different cell types, can be usefully characterized starting with a whole-organ approach. The analysis, based on standard pathways and in situ hybridization studies, begins to suggest correlated gene functions for distinct processes taking place at the same time in newborn ovary. In particular, we detected several keratin genes downregulated in adult ovary as well as several known and previously unreported transcripts specific for primordial and small growing follicles. They included examples expressed in oocytes (*c-kit*, *RLTR10*), in granulosa cells (*Wnt4*), or in both (*cyclin D3*). In addition, coordinated changes in developmentally relevant gene pathways permitted us to

- identify several genes enriched in oocytes at different follicle stages (Table 3), including a member of the Notch pathway (*Dlx2*) and a transcription factor with oocyte-specific functions in the fruit fly (*BicD2*) (Figs. 6C–F);

- suggest candidate somatic genes with differential expression (Table 4), including some that are possibly associated with the morphogenetic mechanisms involving migrating stromal and endothelial cells during newborn ovary septation (Rajah et al., 1992) (notably *Podxl*, Fig. 4; and PDGFR-β, Figs. 5E–F); and
- reveal a striking enrichment of retrotransposon-related sequences in the newborn ovary (*RLTR10*, Figs. 4 and 6A–B, and WEB Fig. 2).

Novel noncoding repeat family

The new, nonautonomous repeat family represented by H3146E11 is highly expressed in the oocytes of primordial and primary follicles, and at lower levels in granulosa cells of multilayered follicles. The probe likely detects a large fraction of the many loci that show extensive similarity to H3146E11 (see Results). It is therefore striking that the observed hybridization signal is highly tissue and stage specific. The data indicate that this does not reflect massive expression of one or a few loci, but rather coexpression of many members of the gene family in newborn ovary.

Though their genomic distribution and context have not been analyzed in detail, some other transposons have been reported to show gonad-specific expression. LINE-1s, for example, are expressed in developing spermatogonia and ovarian cells but not in peri-implantation embryos (Trelogan and Martin, 1995); and IAP sequences and some HERV-K elements are expressed in sperm precursors but not in the ovary (Casau et al., 1999; Dupressoir and Heidmann, 1996; Ono, 1986). Specificity for ovary has also been reported for some VL30 (Schiff et al., 1991), partially characterized LTR cDNAs (Goto et al., 1999, 2002), and one rat IAP ("IAP-LE", for IAP-like element) expressed in granulosa cells of primary follicles (Graham et al., 2000); its LTR elements show 74% identity to the class reported here and may represent a related class.

There has been a tendency to rationalize the expression of such sequences as simply indicating "permissiveness" of transcription in gonadal tissues, but the results rather indicate that transcription is not "general" or "promiscuous", but can be restricted to specific subgroups or classes of sequence, gonadal types, and even stages of development. One possibility is that such retrosequences could be merely carried along as "selfish RNA". Alternatively, elements like those joined to downstream exons of genes (WEB Fig. 2) may operate to regulate development by inhibiting transcription, as reported for other retrotransposon elements (reviewed by Brosius, 1999).

Suggestively, one recently described human-specific LTR retrotransposon family expressed in oocytes (though at all stages rather than primarily early on) was correctly transcribed in oocytes in transgenic fish (Pi et al., 2004). Dissection of the determinants of repeat sequence tissue specificity may help to explain the regulation of such RNAs and of coding RNAs with similar promoter elements.

Prospects for further analysis

Some genes represented on the microarrays and known to be differentially expressed by published reports (e.g., H3005C08, Laminin α5; Frojdman et al., 1999) or by inference from EST analysis and confirmed by us (data not shown) were not scored as significant. Also, other well-characterized factors known to be specifically required for newborn ovary histogenesis (*NGFR*, *TrkB*, *NGF*, *NT4*, e.g., Dissen et al., 2001; Spears et al., 2003) were absent from the 15K array. These deficiencies should be remedied as genes expressed below a threshold value, which have been conservatively excluded, are accurately quantitated with novel technologies (reviewed by Lam and Renil, 2002), and microarrays are used that incorporate more complete gene cohorts.

Thus far, only some of the genes differentially expressed in microarray analyses could be assigned to specific cell types or processes. The assignments could be augmented as EST libraries add more differential expression data. In addition, further fractionation of developmental steps of the immature ovary could be used to assess expression at different follicle stages (similar to the analysis of stages in spermatogenesis in the prepubertal testis; Fujii et al., 2002).

As gene annotations expand and smaller changes in expression can be dependably determined, analyses can be extended in important directions. First, apparent follicle formation from embryonic stem cells in vitro (Hubner et al., 2003) may offer a model for comparative studies of development based purely on derived oocytes and granulosa cells. Second, recent evidence suggests a possibly comparable function of germ cells with stem cell-like characteristics in regeneration of part of the follicle pool postnatally (Johnson et al., 2004). It should be possible to determine how much they contribute to follicle dynamics and how much they depend on a "niche" that regulates growth. Finally, for many instances of haploinsufficiency (like *FOXL2* lesions, in which mutation of one allele can result in early menopause; Crisponi et al., 2001), functionally relevant target genes and networks can be determined for possible intervention in reproductive disorders.

Acknowledgments

We thank Drs. Kevin Becker for preparation of cDNA membranes; Steven Cole (UCLA) for help with microarray analysis; Tetsuya Tanaka for help with microarray techniques; Manuela Uda for help with molecular biology techniques; Yulan Piao for amplified cDNA probes; Carole Stagg and Shakib Omari for help with sample collection and processing; and Dawood Dudekula and Yong Qian for computer expertise and submissions.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ydbio.2004.11.029.

References

Appert, A., Fridmacher, V., Locquet, O., Magre, S., 1998. Patterns of keratins 8, 18 and 19 during gonadal differentiation in the mouse: sex- and time-dependent expression of keratin 19. Differentiation 63, 273–284.

Audic, S., Claverie, J.M., 1997. The significance of digital gene expression profiles. Genome Res. 7, 986–995.

Brosius, J., 1999. RNAs from all categories generate retrosequences that may be expected as novel genes or regulatory elements. Gene 238, 115–134.

Burke, P.S., Wolgemuth, D.J., 1992. Zfp-37, a new murine zinc finger encoding gene, is expressed in a developmentally regulated pattern in the male germ line. Nucleic Acids Res. 20, 2827–2834.

Burns, K.H., Yan, C., Kumar, T.R., Matzuk, M.M., 2001. Analysis of ovarian gene expression in follicle-stimulating hormone β knockout mice. Endocrinology 142, 2742–2751.

Byskov, A.G., Guoliang, X., Andersen, C.Y., 1997. The cortex-medulla oocyte growth pattern is organized during fetal life: an in-vitro study of the mouse ovary. Mol. Hum. Reprod. 3, 795–800.

Canipari, R., 2000. Oocyte-granulosa cell interactions. Hum. Reprod. Update 6, 279–289.

Casau, A.E., Vaughan, J.E., Lozano, G., Levine, A.J., 1999. Germ cell expression of an isolated human endogenous retroviral long terminal repeat of the HERV-K/HTDV family in transgenic mice. J. Virol. 73, 9976–9983.

Carninci, P., Waki, K., Shiraki, T., Konno, H., Shibata, K., Itoh, M., Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Sugahara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Watahiki, A., Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K., Pavan, W., Aidinis, V., Nakagawara, A., Held, W.A., Iwata, H., Kono, T., Nakauchi, H., Lyons, P., Wells, C., Hume, D.A., Fagiolini, M., Hensch, T.K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawai, J., Hayashizaki, Y., 2003. Targeting a complex transcriptome: the construction of the mouse full-length cDNA encyclopedia. Genome Res. 13, 1273–1289.

Castrillon, D.H., Miao, L., Kollipara, R., Horner, J.W., DePinho, R.A., 2003. Suppression of ovarian follicle activation in mice by the transcription factor Foxo3a. Science 301, 215–218.

Caubit, X., Core, N., Boned, A., Kerridge, S., Djabali, M., Fasano, L., 2000. Vertebrate orthologues of the Drosophila region-specific patterning gene teashirt. Mech. Dev. 91, 445–448.

Chin, K.-V., Seifer, D.B., Feng, B., Lin, Y., Shih, W.-C., 2002. DNA microarray analysis of the expression profiles of luteinized granulosa cells as a function of ovarian reserve. Fertil. Steril. 77, 1214–1218.

Cole, S.W., Galic, Z., Zack, J.A., 2003. Controlling false-negative errors in microarray differential expression analysis: a PRIM approach. Bioinformatics 19, 1808–1816.

Crisponi, L., Deiana, M., Loi, A., Chiappe, F., Uda, M., Amati, P., Bisceglia, L., Zelante, L., Nagaraja, R., Porcu, S., et al., 2001. The putative forkhead transcription factor FOXL2 is mutated in belopharophimosis/ptosis/epicanthus inversus syndrome. Nat. Genet. 27, 159–166.

Cui, C.Y., Durmowicz, M., Tanaka, T.S., Hartung, A.J., Tezuka, T., Hashimoto, K., Ko, M.S., Srivastava, A.K., Schlessinger, D., 2002. EDA targets revealed by skin gene expression profiles of wild-type, Tabby and Tabby EDA-A1 transgenic mice. Hum. Mol. Genet. 11, 1763–1773.

Dean, J., 2002. Oocyte-specific genes regulate follicle formation, fertility and early mouse development. J. Reprod. Immun. 53, 171–180.

de Gooyer, T.E., Skinner, S.L., Wlodek, M.E., Kelly, D.J., Wilkinson-Berka, J.L., 2004. Angiotensin II influences ovarian follicle development in the transgenic (mRen-2)27 and Sprague–Dawley rat. J. Endocrinol. 180, 311–324.

Dissen, G.A., Hirshfield, A.N., Malamed, S., Ojeda, S.R., 1995. Expression of neurotrophins and their receptors in the mammalian ovary is developmentally regulated: changes at the time of folliculogenesis. Endocrinology 136, 4681–4692.

Dissen, G.A., Romero, C., Hirshfield, A.N., Ojeda, S.R., 2001. Nerve growth factor is required for early follicular development in the mammalian ovary. Endocrinology 142, 2078–2086.

Dong, J., Albertini, D.F., Nishimori, K., Kumar, T.R., Lu, N., Matzuk, M.M., 1996. Growth differentiation factor-9 is required during early ovarian folliculogenesis. Nature 383, 485–486.

Doyonnas, R., Kershaw, D.B., Duhme, C., Merkens, H., Chelliah, S., Graf, T., McNagny, K.M., 2001. Anuria, omphalocele, and perinatal lethality in mice lacking the CD34-related protein podocalyxin. J. Exp. Med. 194, 13–27.

Drummond, A.E., Dyson, M., Mercer, J.E., Findlay, J.K., 1996. Differential responses of post-natal rat ovarian cells to FSH and activin. Mol. Cell. Endocrinol. 122, 21–32.

Dupressoir, A., Heidmann, T., 1996. Germ line-specific expression of intracisternal A-particle retrotransposons in transgenic mice. Mol. Cell. Biol. 8, 4495–4503.

Elvin, J.A., Matzuk, M.M., 1998. Mouse models of ovarian failure. Rev. Reprod. 3, 183–195.

Elvin, J.A., Yan, C., Wang, P., Nishimori, K., Matzuk, M.M., 1999. Molecular characterization of the follicle defects in the growth differentiation factor 9-deficient ovary. Mol. Endocrinol. 13, 1018–1034.

Flaggs, G., Plug, A.W., Dunks, K.M., Mundt, K.E., Ford, J.C., Quiggle, M.R.E., Taylor, E.M., Westphal, C.H., Ashley, T., Hoekstra, M.F., Carr, A.M., 1997. Atm-dependent interactions of a mammalian Chk1 homolog with meiotic chromosomes. Curr. Biol. 7, 977–986.

Frojdman, K., Miner, J.H., Sanes, J.R., Pelliniemi, L.J., Virtanen, I., 1999. Sex-specific localization of laminin alpha 5 chain in the differentiating rat testis and ovary. Differentiation 64, 151–159.

Fujii, T., Tamura, K., Masai, K., Tanaka, H., Nishimune, Y., Nojima, H., 2002. Use of stepwise subtraction to comprehensively isolate mouse genes whose transcription is up-regulated during spermiogenesis. EMBO Rep. 3, 367–372.

Goto, T., Adjaye, J., Rodeck, C.H., Monk, M., 1999. Identification of genes expressed in human primordial germ cells at the time of entry of the female germ line into meiosis. Mol. Hum. Reprod. 5, 851–860.

Goto, T., Jones, G.M., Lolatgis, N., Pera, M.F., Trounson, A.O., Monk, M., 2002. Identification and characterisation of known and novel transcripts expressed during the final stages of human oocyte maturation. Mol. Reprod. Dev. 62, 13–28.

Graham, K.M., Ko, C., Park, K.S., Sarge, K., Park-Sarge, O.K., 2000. Expression of an intracisternal A-particle-like element in rat ovary. Biochem. Biophys. Res. Commun. 278, 48–57.

Grimmond, S., Van Hateren, N., Siggers, P., Arkell, R., Larder, R., Soares, M.B., de Fatima Bonaldo, M., Smith, L., Tymowska-Lalanne, Z., Wells, C., et al., 2000. Sexually dimorphic expression of protease nexin-1 and vanin-1 in the developing mouse gonad prior to overt differentiation suggests a role in mammalian development. Hum. Mol. Genet. 9, 1553–1560.

Hara, T., Nakano, Y., Tanaka, M., Tamura, K., Sekiguchi, T., Minehata, K., Copeland, N.G., Jenkins, N.A., Okabe, M., Kogo, H., Mukouyama, Y., Miyajima, A., 1999. Identification of podocalyxin-like protein 1 as a novel cell surface marker for hemangioblasts in the murine aorta-gonad-mesonephros region. Immunity 11, 567–578.

Hergersberg, M., Matsuo, K., Gassmann, M., Schaffner, W., Luscher, B., Rulicke, T., Aguzzi, A., 1995. Tissue-specific expression of a FMR1/β-galactosidase fusion gene in transgenic mice. Hum. Mol. Genet. 4, 359–366.

Ho, M., Yang, E., Matcuk, G., Deng, D., Sampas, N., Tsalenko, A., Tabibiazar, R., Zhang, Y., Chen, M., Talbi, S., Ho, Y.D., Wang, J., Tsao, P.S., Ben-Dor, A., Yakhini, Z., Bruhn, L., Quertermous, T., 2003. Identification of endothelial cell genes by combined database mining and microarray analysis. Physiol. Genomics 13, 249–262.

Hoogenraad, C.C., Akhmanova, A., Howell, S.A., Dortland, B.R., De Zeeuw, C.I., Willemsen, R., Visser, P., Grosveld, F., Galjart, N., 2001. Mammalian Golgi associated Bicaudal-D2 functions in the dynein–dynactin pathway by interacting with these complexes. EMBO J. 20, 4041–4054.

Hubner, K., Fuhrmann, G., Christenson, L.K., Kehler, J., Reinbold, R., De La Fuente, R., Wood, J., Strauss III, J.F., Boiani, M., Scholer, H.R., 2003. Derivation of oocytes from mouse embryonic stem cells. Science 300, 1251–1256.

Huminiecki, L., Bicknell, R., 2000. In silico cloning of novel endothelial-specific genes. Genome Res. 10, 1796–1806.

Johnson, J., Canning, J., Kaneko, T., Pru, J.K., Tilly, J.L., 2004. Germline stem cells and follicular renewal in the postnatal mammalian ovary. Nature 428, 145–150.

Kakuchi, J., Ichiki, T., Kiyama, S., Hogan, B.L., Fogo, A., Inagami, T., Ichikawa, I., 1995. Developmental expression of renal angiotensin II receptor genes in the mouse. Kidney Int. 47, 140–147.

Kishi, N., Tang, Z., Maeda, Y., Hirai, A., Mo, R., Ito, M., Suzuki, S., Nakao, K., Kinoshita, T., Kadesch, T., Hui, C., Artavanis-Tsakonas, S., Okano, H., Matsuno, K., 2001. Murine homologs of deltex define a novel gene family involved in vertebrate Notch signaling and neurogenesis. Int. J. Dev. Neurosci. 9, 21–35.

Koopmann, P., Bullejos, M., Loffler, K., Bowles, J., 2002. Expression-based strategies for discovery of genes involved in testis and ovary development. Novartis Found. Symp. 244, 240–249.

Kroll, T.C., Wolfl, S., 2002. Ranking: a closer look on globalisation methods for normalisation of gene expression arrays. Nucleic Acids Res. 30, e50.

Kuff, E.L., Lueders, K.K., 1988. The intracisternal A-particle gene family: structure and functional aspects. Adv. Cancer Res. 51, 183–276.

Lam, K.S., Renil, M., 2002. From combinatorial chemistry to chemical microarray. Curr. Opin. Chem. Biol. 6, 353–358.

Liang, L.-F., Soyal, S.M., Dean, J., 1997. FIG-α, a germ cell specific transcription factor involved in the coordinate expression of the zona pellucida genes. Development 124, 4939–4947.

Linder, B., Jones, L.K., Chaplin, T., Mohd-Sarip, A., Heinlein, U.A., Young, B.D., Saha, V., 1998. Expression pattern and cellular distribution of the murine homologue of AF10. Biochim. Biophys. Acta 1443, 285–296.

Liu, H.-C., He, Z., Rosenwaks, Z., 2001. Application of complementary DNA microarray (DNA chip) technology in the study of gene expression profiles during folliculogenesis. Fertil. Steril. 75, 947–955.

Liu, C., Zhang, L., Shao, Z.M., Beatty, P., Sartippour, M., Lane, T.F., Barsky, S.H., Livingston, E., Nguyen, M., 2002. Identification of a novel endothelial-derived gene EG-1. Biochem. Biophys. Res. Commun. 290, 602–612.

McLaren, A., 2000. Germ and somatic cell lineages in the developing gonad. Mol. Cell. Endocrinol. 163, 3–9.

Magnaldo, T., Fowlis, D., Darmon, M., 1998. Galectin-7, a marker of all types of stratified epithelia. Differentiation 63, 159–168.

Manova, K., Nocka, K., Besmer, P., Bachvarova, R.F., 1990. Gonadal expression of c-kit encoded at the W locus of the mouse. Development 110, 1057–1069.

Matsui, Y., Zsebo, K.M., Hogan, B.L., 1990. Embryonic expression of a haematopoietic growth factor encoded by the Sl locus and the ligand for c-kit. Nature 347, 667–669.

Mazarakis, N., Michalovich, D., Karis, A., Grosveld, F., Galjart, N., 1996. Zfp-37 is a member of the KRAB zinc finger gene family and is expressed in neurons of the developing and adult CNS. Genomics 33, 247–257.

Meagher, M.J., Schumacher, J.M., Lee, K., Holdcraft, R.W., Edelhoff, S., Disteche, C., Braun, R.E., 1999. Identification of ZFR, an ancient and highly conserved murine chromosome-associated zinc finger protein. Gene 228, 197–211.

Nishimura, H., Yerkes, E., Hohenfellner, K., Miyazaki, Y., Ma, J., Hunley, T.E., Yoshida, H., Ichiki, T., Threadgill, D., Phillips III, J.A., Hogan, B.M., Fogo, A., Brock III, J.W., Inagami, T., Ichikawa, I., 1999. Role of the angiotensin type 2 receptor gene in congenital anomalies of the kidney and urinary tract, CAKUT, of mice and men. Mol. Cell 3, 1–10.

Ohno, S., Smith, J.B., 1964. Role of fetal follicular cells in meiosis of mammalian oocytes. Cytogenetics 13, 324–333.

Ono, M., 1986. Molecular cloning and long terminal repeat sequences of human endogenous retrovirus genes related to types A and B retrovirus genes. J. Virol. 3, 937–944.

Park, T., Yi, S.G., Kang, S.H., Lee, S., Lee, Y.S., Simon, R., 2003. Evaluation of normalization methods for microarray data. BMC Bioinformatics 4, 33.

Patel, K., Connolly, D.J., Amthor, H., Nose, K., Cooke, J., 1996. Cloning and early dorsal axial expression of Flik, a chick follistatin-related gene: evidence for involvement in dorsalization/neural induction. Dev. Biol. 178, 327–342.

Pellegrini, M., Pantano, S., Lucchini, F., Fumi, M., Forabosco, A., 1997. Emx2 developmental expression in the primordia of the reproductive and excretory systems. Anat. Embryol. 196, 427–433.

Perera, E.M., Martin, H., Seeherunvong, T., Kos, L., Hughes, I.A., Hawkins, J.R., Berkovitz, G.D., 2001. Tescalcin, a novel gene encoding a putative EF-hand $Ca^{(2+)}$-binding protein, Col9a3, and renin are expressed in the mouse testis during the early stages of gonadal differentiation. Endocrinology 142, 455–463.

Peters, H., 1969. The development of the mouse ovary from birth to maturity. Acta Endocrinol. (Copenh.) 62, 98–116.

Pi, W., Yang, Z., Wang, J., Ruan, L., Yu, X., Ling, J., Krantz, S., Isales, C., Conway, S.J., Lin, S., Tuan, D., 2004. The LTR enhancer of ERV-9 human endogenous retrovirus is active in oocytes and progenitor cells in transgenic zebrafish and humans. Proc. Natl. Acad. Sci. U. S. A. 101, 805–810.

Pointud, J.C., Mengus, G., Brancorsini, S., Monaco, L., Parvinen, M., Sassone-Corsi, P., Davidson, I., 2003. The intracellular localisation of TAF7L, a paralogue of transcription factor TFIID subunit TAF7, is developmentally regulated during male germ-cell differentiation. J. Cell Sci. 116, 1847–1858.

Pru, J.K., Tilly, J.L., 2001. Programmed cell death in the ovary: insights and future prospects using genetic technologies. Mol. Endocrinol. 15, 845–853.

Puglianiello, A., Campagnolo, L., Farini, D., Cipollone, D., Russo, M.A., Siracusa, G., 2004. Expression and role of PDGF-BB and PDGFR-β during testis morphogenesis in the mouse embryo. J. Cell Sci. 117, 1151–1160.

Rajah, R., Glaser, E.M., Hirshfield, A.N., 1992. The changing architecture of the neonatal rat ovary during histogenesis. Dev. Dyn. 194, 177–192.

Rajkovic, A., Matzuk, M.M., 2002. Functional analysis of oocyte-expressed genes using transgenic models. Mol. Cell. Endocrinol. 187, 5–9.

Rajkovic, A., Pangas, S.A., Ballow, D., Suzumori, N., Matzuk, M.M., 2004. NOBX deficiency disrupts early folliculogenesis and oocyte-specific gene expression. Science 305, 1157–1159.

Reynaud, K., Driancourt, M.A., 2000. Oocyte attrition. Mol. Cell. Endocrinol. 163, 101–108.

Rodgers, R.J., Irving-Rodgers, H.F., van Wezel, I.L., Krupa, M., Lavranos, T.C., 2001. Dynamics of the membrana granulosa during expansion of the ovarian follicular antrum. Mol. Cell. Endocrinol. 171, 41–48.

Ruggiu, M., Speed, R., Taggart, M., McKay, S.J., Kilanowski, F., Saunders, P., Dorin, J., Cooke, H.J., 1997. The mouse Dazla gene encodes a cytoplasmic protein essential for gametogenesis. Nature 389, 73–77.

Schiff, R., Itin, A., Keshet, E., 1991. Transcriptional activation of mouse retrotransposons in vivo: specific expression in steroidogenic cells in response to trophic hormones. Genes Dev. 5, 521–532.

Schmidt, D., Ovitt, C.E., Anlag, K., Fehsenfeld, S., Gredsted, L., Treier, A.C., Treier, M., 2004. The murine winged-helix transcription factor Foxl2 is required for granulosa cell differentiation and ovary maintenance. Development 131, 933–942.

Sforza, C., Ferrario, V.F., De Pol, A., Marzona, L., Forni, M., Forabosco, A., 1993. Morphometric study of the human ovary during compartmentalization. Anat. Rec. 236, 626–634.

Sharov, A.A., Piao, Y., Matoba, R., Dudekula, D.B., Qian, Y., VanBuren, V., Falco, G., Martin, P.R., Stagg, C.A., Bassey, U.C., Wang, Y., Carter, M.G., Hamatani, T., Aiba, K., Akutsu, H., Sharova, L., Tanaka, T.S., Kimber, W.L., Yoshikawa, T., Jaradat, S.A., Pantano, S., Nagaraja, R., Boheler, K.R., Taub, D., Hodes, R.J., Longo, D.L., Schlessinger, D., Keller, J., Klotz, E., Kelsoe, G., Umezawa, A., Vescovi, A.L., Rossant, J., Kunath, T., Hogan, B.L., Curci, A., D'Urso, M., Kelso, J., Hide, W., Ko, M.S., 2003. Transcriptome analysis of mouse stem cells and early embryos. PLoS Biol. 1, E74.

Shi, F., LaPolt, P.S., 2003. Relationship between Foxo1 protein levels and follicular development, atresia, and luteinization in the rat ovary. J. Endocrinol. 179, 195–203.

Shibanuma, M., Mashimo, J., Mita, A., Kuroki, T., Nose, K., 1993. Cloning from a mouse osteoblastic cell line of a set of transforming-growth-factor-β 1-regulated genes, one of which seems to encode follistatin-related polypeptide. Eur. J. Biochem. 217, 13–19.

Shibata, H., Kurihara, I., Kobayashi, S., Yokota, K., Suda, N., Saito, I., Saruta, T., 2003. Regulation of differential COUP-TF-coregulator interactions in adrenal cortical steroidogenesis. J. Steroid Biochem. Mol. Biol. 85, 449–456.

Sidorov, I.A., Hasack, D.A., Gee, D., Yang, J., Cam, M.C., Lempicki, R.A., Dimitrov, D.S., 2002. Oligonucleotide microarray data distribution and normalization. Inf. Sci. 146, 67–73.

Soyal, S.M., Amieh, A., Dean, J., 2000. FIG-α, a germ cell-specific transcription factor required for ovarian follicle formation. Development 127, 4645–4654.

Spears, N., Molinek, M.D., Robinson, L.L., Fulton, N., Cameron, H., Shimoda, K., Telfer, E.E., Anderson, R.A., Price, D.J., 2003. The role of neurotrophin receptors in female germ-cell survival in mouse and human. Development 130, 5481–5491.

Storey, J.D., Tibshirani, R., 2003. Statistical significance for genomewide studies. Proc. Natl. Acad. Sci. U. S. A. 100, 9440–9445.

Su, A.I., Cooke, M.P., Ching, K.A., Hakak, Y., Walker, J.R., Wiltshire, T., Orth, A.P., Vega, R.G., Sapinoso, L.M., Mogrich, A., Patapoutian, A., Hampton, G.M., Schultz, P.G., Hogenesch, J.B., 2002. Large-scale analysis of the human and mouse transcriptomes. Proc. Natl. Acad. Sci. U. S. A. 99, 4465–4470.

Takasaki, N., McIsaac, R., Dean, J., 2000. Gpbox (Psx2), a homeobox gene preferentially expressed in female germ cells at the onset of sexual dimorphism in mice. Dev. Biol. 223, 181–193.

Tanaka, T.S., Jaradat, S.A., Lim, M.K., Kargul, G.J., Wang, X., Grahovac, M.J., Pantano, S., Sano, Y., Piao, Y., Nagaraja, R., et al., 2000. Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray. Proc. Natl. Acad. Sci. U. S. A. 97, 9127–9132.

Tong, Z.B., Nelson, L.M., 2000. Mater encodes a maternal protein in mice with a leucine-rich repeat domain homologous to porcine ribonuclease inhibitor. Mamm. Genome 11, 281–287.

Trelogan, S.A., Martin, S.L., 1995. Tightly regulated, developmentally specific expression of the first open reading frame from LINE-1 during mouse embryogenesis. Proc. Natl. Acad. Sci. U. S. A. 92, 1520–1524.

Tusher, V.G., Tibshirani, R., Chu, G., 2001. Significance analysis of microarrays applied to the ionizing radiation response. Proc. Natl. Acad. Sci. U. S. A. 98, 5116–5121.

Uda, M., Ottolenghi, C., Crisponi, L., Garcia, J.E., Deiana, M., Kimber, W., Forabosco, A., Cao, A., Schlessinger, D., Pilia, G., 2004. Foxl2 disruption causes mouse ovarian failure by pervasive blockage of follicle development. Hum. Mol. Genet. 13, 1171–1181.

Vainio, S., Heikkila, M., Kispert, A., Chin, N., McMahon, A.P., 1999. Female development in mammals is regulated by Wnt-4 signaling. Nature 397, 405–409.

Vanani, S., Elvin, J.A., Uan, C., DeMayo, J., DeMayo, J., Horton, H.F., Byrne, M.C., Matzuk, M.M., 2002. Knockout of Pentraxin 3, a downstream target of growth differentiation factor-9, causes female subfertility. Mol. Endocrinol. 16, 1154–1167.

Vitt, U.A., McGee, E.A., Hayashi, M., Hsueh, A.J., 2000. In vivo treatment with GDF-9 stimulates primordial and primary follicle progression and theca cell marker CYP17 in ovaries of immature rats. Endocrinology 141, 3814–3820.

Wertz, K., Herrmann, B.G., 2000. Large-scale screen for genes involved in gonad development. Mech. Dev. 98, 51–70.

Wilkinson, D.G., Nieto, M.A., 1993. Detection of messenger RNA by in situ hybridization to tissue sections and whole mounts. Methods Enzymol. 225, 361–373.

Yamaguchi, F., Morrison, R.S., Gonatas, N.K., Takahashi, H., Sugisaki, Y., Teramoto, A., 2003. Identification of MG-160, a FGF binding medial Golgi sialoglycoprotein, in brain tumors: an index of malignancy in astrocytomas. Int. J. Oncol. 22, 1045–1049.

Yan, C., Wang, P., DeMayo, J., Demayo, F.J., Elvin, J.A., Carino, C., Prasad, S.V., Skinner, S.S., Dunbar, B.S., Dube, J.L., Celeste, A.J., Matzuk, M.M., 2001. Synergistic roles of bone morphogenetic protein 15 and growth differentiation factor 9 in ovarian function. Mol. Endocrinol. 15, 854–866.

Zhang, Q., Wang, X., Wolgemuth, D.J., 1999. Developmentally regulated expression of cyclin D3 and its potential in vivo interacting proteins during murine gametogenesis. Endocrinology 140, 2790–2800.

Web site references

http://www.lgsun.grc.nia.nih.gov/microarray/herrera, with all Supplementary Material
http://lgsun.grc.nia.nih.gov/geneindex/, with data on the membrane arrays.
http://www.ncbi.nlm.nih.gov, with sites therein for OMIM, etc.
http://fantom2.gsc.riken.go.jp
http://bioinfo.weizmann.ac.il/cards/index.shtml
http://www.girinst.org/Repbase_Reports
http://ftp.genome.washington.edu/cgi-bin/RepeatMasker
http://www.ensembl.org/mus_musculus
http://blast.wustl.edu
http://microarray.crump.ucla.edu/focus